

# Graph Lambda Theories <sup>\*</sup>

Antonio Bucciarelli

Antonino Salibra <sup>†</sup>

## Abstract

A longstanding open problem in lambda calculus is whether there exist continuous models of the untyped lambda calculus whose theory is exactly the beta-theory or the least sensible  $\lambda$ -theory  $\mathcal{H}$  (generated by equating all the unsolvable terms). A related question, raised recently by C. Berline, is whether, given a class of lambda models, there are a minimal  $\lambda$ -theory and a minimal sensible  $\lambda$ -theory represented by it. In this paper, we give a positive answer to this question for the class of graph models à la Plotkin-Scott-Engeler. In particular, we build two graph models whose theories are respectively the set of equations satisfied in any graph model and in any sensible graph model. We conjecture that the least sensible graph theory, where “graph theory” means “ $\lambda$ -theory of a graph model”, is equal to  $\mathcal{H}$ , while in one of the main results of the paper we show the non-existence of a graph model whose equational theory is exactly the beta-theory (this result negatively answers Question 1 in [7, Section 6.2] for the restricted class of graph models).

Another related question is whether, given a class of lambda models, there is a maximal sensible  $\lambda$ -theory represented by it. In the main result of the paper we characterize the greatest sensible graph theory as the  $\lambda$ -theory  $\mathcal{B}$  generated by equating  $\lambda$ -terms with the same Böhm tree. This result is a consequence of the main technical theorem of the paper: all the equations between solvable  $\lambda$ -terms, which have different Böhm trees, fail in every sensible graph model. A further result of the paper is the existence of a continuum of different sensible graph theories strictly included in  $\mathcal{B}$  (this result positively answers Question 2 in [7, Section 6.3]).

**Keywords.** *Lambda calculus, lambda theories, graph models, minimum graph theory, maximum graph theory, beta-theory.*

## 1 Introduction

The untyped lambda calculus was introduced around 1930 by Church [14, 15] as part of an investigation in the formal foundations of mathematics and logic. Although lambda calculus is a very basic language, it is sufficient to express all computable functions. The process

---

<sup>\*</sup>This article is a revised and expanded version of two papers of the authors which appeared respectively in the Proceedings of the 28th International Symposium on Mathematical Foundations of Computer Science (MFCS’03) [12], and in the Proceedings of the 19th Annual IEEE Symposium on Logic in Computer Science (LICS’04) [13].

<sup>†</sup>Work partially supported by the Equipe PPS of the University Paris 7-Denis Diderot, and by MIUR Cofin’04 FOLLIA Project.

of application and evaluation reflects the computational behavior of many modern functional programming languages, which explains the interest in the lambda calculus among computer scientists.

Lambda theories are equational extensions of the untyped lambda calculus closed under derivation. They arise by syntactical or semantic considerations. Indeed, a  $\lambda$ -theory may correspond to a possible operational (observational) semantics of the lambda calculus, as well as it may be induced by a model of lambda calculus through the kernel congruence relation of the interpretation function. Although researchers have mainly focused their interest on a limited number of them, the class of  $\lambda$ -theories constitutes a very rich and complex structure (see e.g. [4, 7]). Syntactical techniques are usually difficult to use in the study of  $\lambda$ -theories. Therefore, semantic methods have been extensively investigated.

Topology is at the center of the known approaches to giving models of the untyped lambda calculus. The first model, found by Scott in 1969 in the category of complete lattices and Scott continuous functions, was successfully used to show that all unsolvable  $\lambda$ -terms can be consistently equated. After Scott, a large number of mathematical models for lambda calculus, arising from syntax-free constructions, have been introduced in various categories of domains and were classified into semantics according to the nature of their representable functions, see e.g. [1, 4, 7, 26]. Scott's continuous semantics [29] is given in the category whose objects are complete partial orders and morphisms are Scott continuous functions. The stable semantics (Berry [10]) and the strongly stable semantics (Bucciarelli-Ehrhard [11]) are a strengthening of the continuous semantics, introduced to capture the notion of "sequential" Scott continuous function. All these semantics are structurally and equationally rich in the sense that it is possible to build up  $2^{\aleph_0}$  models in each of them inducing pairwise distinct  $\lambda$ -theories [23, 24]. Nevertheless, the above denotational semantics are *equationally incomplete*: they do not match all possible operational semantics of lambda calculus. The problem of the equational incompleteness was positively solved by Honsell-Ronchi della Rocca [21] for the continuous semantics, and by Bastonero-Gouy [20, 6] for the stable semantics. Salibra [27, 28] has recently shown in a uniform way that all semantics, which involve monotonicity with respect to some partial order and have a bottom element, fail to induce a continuum of  $\lambda$ -theories. From this it follows the incompleteness of the strongly stable semantics, which had been conjectured by Bastonero-Gouy [6] and by Berline [7].

If a semantics is incomplete, then there exists a  $\lambda$ -theory  $T$  that is not induced by any model in the semantics. In such a case we say that the semantics *omits* the  $\lambda$ -theory  $T$ . More generally, a semantics *omits* (*forces*, respectively) an equation if the equation fails (holds) in every model of the semantics. The set of equations forced by a semantics  $\mathcal{C}$  constitutes a  $\lambda$ -theory. It is the minimal (with respect to the inclusion order)  $\lambda$ -theory of  $\mathcal{C}$  if it is induced by a model of  $\mathcal{C}$ .

The following natural questions arises (see Berline [7]): given a class  $\mathcal{C}$  of models of lambda calculus,

1. Is there a minimal  $\lambda$ -theory represented by  $\mathcal{C}$ ?
2. Is there a minimal sensible (i.e., equating all unsolvable  $\lambda$ -terms)  $\lambda$ -theory represented

by  $\mathcal{C}$ ?

Di Gianantonio et al. [19] have shown that the above question (1) admits a positive answer for Scott's continuous semantics, at least if we restrict to extensional models. However, the proofs of [19] use logical relations, and since logical relations do not allow to distinguish terms with the same applicative behavior, the proofs do not carry out to non-extensional models.

In this paper we show that both question (1) and question (2) admit a positive answer for the *graph semantics*, that is, the semantics of lambda calculus given in terms of *graph models*. These models, isolated in the seventies by Scott and Engeler [4] within the continuous semantics, have been proved useful for giving proofs of consistency of extensions of lambda calculus and for studying operational features of lambda calculus. For example, the simplest graph model, namely Engeler's model, has been used by Berline [7] to give concise proofs of the head-normalization theorem and of the left-normalization theorem of lambda calculus, while a semantical proof of the "easiness" of  $(\lambda x.xx)(\lambda x.xx)$  was obtained by Baeten and Boerboom in [3]. It is well known that the graph semantics is incomplete, since it trivially omits the axiom of extensionality. The main technical device used in the proof of the existence of the least (sensible) graph theory is the notion of *weak product* of graph models. Roughly speaking, the weak product of a family of graph models is a new graph model which is the "canonical completion" of the disjoint union of the models in the family. We show that the theory of a weak product is always semisensible (i.e., it does not equate solvable and unsolvable terms) and it is included in the intersection of the theories of its factors (the inclusion is in general strict). The *least graph theory* (where "graph theory" means " $\lambda$ -theory of a graph model") is the theory of the weak product of the family  $(D_e : e \in I)$ , where  $I$  is the set of equations between  $\lambda$ -terms which fail to hold in some graph model, and  $D_e$  is a fixed graph model not satisfying the equation  $e$ .

Two further questions naturally arise: what equations between  $\lambda$ -terms belong to the minimal graph theory? And to the minimal sensible one? The answer to the second difficult question is still unknown; we conjecture that the  $\lambda$ -theory  $\mathcal{H}$ , generated by equating all unsolvable  $\lambda$ -terms, is the least sensible graph theory. The first question is related to a longstanding open problem in lambda calculus, asking whether there exists a non-syntactic model whose equational theory is equal to the least  $\lambda$ -theory  $\lambda\beta$ . In this paper we show that this model cannot be found within graph semantics (this result negatively answers Question 1 in [7, Section 6.2] for the restricted class of graph models). From this result it follows that the minimal graph theory is not equal to  $\lambda\beta$ , so that graph semantics forces equations between non- $\beta$ -equivalent  $\lambda$ -terms. In this paper we provide an example of an equation of this kind.

The set of all sensible  $\lambda$ -theories constitutes a bounded lattice. The least sensible  $\lambda$ -theory is the  $\lambda$ -theory  $\mathcal{H}$  (generated by equating all the unsolvable terms), while the greatest sensible  $\lambda$ -theory is the  $\lambda$ -theory  $\mathcal{H}^*$  (generated by equating terms with the same Böhm tree up to possibly infinite  $\eta$ -equivalence). Kerth has shown in [23] that there exists a continuum of different sensible graph theories. Then it make sense to ask whether there exists a maximal  $\lambda$ -theory represented by graph semantics. In one of the main results of the paper we show that the  $\lambda$ -theory  $\mathcal{B}$  (generated by equating  $\lambda$ -terms with the same Böhm tree) is the greatest

sensible graph theory. This result is a consequence of the main technical theorem of the paper: the graph semantics omits all equations  $M = N$  between  $\lambda$ -terms satisfying the following conditions:

$$\mathcal{H}^* \vdash M = N \text{ and } \mathcal{B} \not\vdash M = N. \quad (1)$$

In other words, the graph semantics omits all equations  $M = N$  between  $\lambda$ -terms which do not have the same Böhm tree, but have the same Böhm tree up to (possibly infinite)  $\eta$ -equivalence.

The following are other consequences of the main result of the paper.

- (i) There exists a continuum of different sensible graph theories strictly included in  $\mathcal{B}$  (this result positively answers Question 2 in [7, Section 6.3]);
- (ii) For every closed term  $P$ , the  $\lambda$ -theory generated by  $\Omega = P$ , where  $\Omega$  is the paradigmatic unsolvable term  $(\lambda x.xx)(\lambda x.xx)$ , contains no equation satisfying condition (1).

The paper is organized as follows. In Section 2 we review the basic definitions of lambda calculus and graph models. In particular, we recall the formal definition of the canonical completion of a partial model. The notion of a weak product of graph models is introduced and studied in Section 3. The proof of the existence of a minimal (sensible) graph theory is presented in Section 4, while in Section 5 it is shown that the least graph theory is not equal to  $\lambda\beta$ . Section 6 is devoted to the characterization of the maximal sensible graph theory. Conclusions and future work are presented in Section 7.

## 2 Preliminaries

To keep this article self-contained, we summarize some definitions and results concerning lambda calculus and graph models that we need in the subsequent part of the paper. With regard to the lambda calculus we follow the notation and terminology of [4].

### 2.1 Lambda calculus

The set  $\Lambda$  of  $\lambda$ -terms over an infinite set of variables is constructed as usual: every variable is a  $\lambda$ -term; if  $M$  and  $N$  are  $\lambda$ -terms, then so are  $(MN)$  and  $\lambda x.M$  for each variable  $x$ .  $\Lambda^o$  denotes the set of closed  $\lambda$ -terms.

The symbol  $\equiv$  denotes syntactic equality. The following are some well-known  $\lambda$ -terms:

$$\Omega \equiv (\lambda x.xx)(\lambda x.xx); \quad \Omega_3 \equiv (\lambda x.xxx)(\lambda x.xxx);$$

$$\mathbf{i} \equiv \lambda x.x; \quad \mathbf{k} \equiv \lambda xy.x; \quad \mathbf{1} \equiv \lambda xy.xy.$$

A *compatible  $\lambda$ -relation*  $T$  is any set of equations between  $\lambda$ -terms that is closed under the following two rules:

- (i) If  $M = N \in T$  and  $P = Q \in T$ , then  $MP = NQ \in T$ ;
- (ii) If  $M = N \in T$  then  $\lambda x.M = \lambda x.N \in T$  for every variable  $x$ .

We will write either  $T \vdash M = N$  or  $M =_T N$  for  $M = N \in T$ .

A  $\lambda$ -theory  $T$  is any compatible  $\lambda$ -relation which is an equivalence relation and includes  $(\alpha)$ - and  $(\beta)$ -conversion. The set of all  $\lambda$ -theories is naturally equipped with a lattice structure, with meet defined as set theoretical intersection. The join of two  $\lambda$ -theories  $T$  and  $S$  is the least equivalence relation including  $T \cup S$ .  $\lambda\beta$  denotes the minimal  $\lambda$ -theory, while  $\lambda\beta\eta$  denotes the minimal extensional  $\lambda$ -theory (axiomatized by  $\mathbf{i} = \mathbf{1}$ ).

Solvable  $\lambda$ -terms can be characterized as follows: a  $\lambda$ -term  $M$  is solvable if, and only if, it has a *head normal form*, that is,  $M =_{\lambda\beta} \lambda x_1 \dots x_n.y M_1 \dots M_k$  for some  $n, k \geq 0$  and  $\lambda$ -terms  $M_1, \dots, M_k$ .  $M \in \Lambda$  is *unsolvable* if it is not solvable.

The  $\lambda$ -theory  $\mathcal{H}$ , generated by equating all unsolvable  $\lambda$ -terms, is consistent by [4, Thm. 16.1.3] and admits a unique maximal consistent extension  $\mathcal{H}^*$  [4, Thm. 16.2.6]. A  $\lambda$ -theory  $T$  is called *sensible* [4, Def. 4.1.7(ii)] if it is consistent and  $\mathcal{H} \subseteq T$ . The set of all sensible  $\lambda$ -theories is naturally equipped with a structure of bounded lattice.  $\mathcal{H}$  is the least sensible  $\lambda$ -theory, while  $\mathcal{H}^*$  is the greatest one.  $\mathcal{H}^*$  is an extensional  $\lambda$ -theory.

A  $\lambda$ -theory is *semisensible* [4, Def. 4.1.7(iii)] if no solvable term is equivalent to an unsolvable term. It is easy to prove that sensible theories are semisensible. It is also possible to characterize semisensible  $\lambda$ -theories as follows: a  $\lambda$ -theory  $T$  is semisensible if, and only if,  $T \subseteq \mathcal{H}^*$  (see Section 16.2 in [4]).

## 2.2 Böhm trees

A  $\lambda$ -term  $M$  is called a *projection term* if  $M \equiv \lambda x_1 \dots x_n.y$  ( $n \geq 0$ ). A *Böhm-like tree* is a finite branching labelled tree, whose inner nodes are labelled by projection terms and leaves either by projection terms or by  $\perp$ .

The Böhm tree  $BT(M)$  of a  $\lambda$ -term  $M$  is a finite or infinite Böhm-like tree. If  $M$  is unsolvable, then  $BT(M) = \perp$ , that is,  $BT(M)$  is a tree with a unique node labelled by  $\perp$ . If  $M$  is solvable and  $\lambda x_1 \dots x_n.y M_1 \dots M_k$  is the principal head normal form of  $M$  [4, Def. 8.3.20] then we have

$$\begin{array}{c}
 BT(M) = \lambda x_1 \dots x_n.y \\
 \swarrow \quad \searrow \\
 BT(M_1) \dots \dots \dots BT(M_k)
 \end{array}$$

The  $\lambda$ -theory  $\mathcal{B}$ , generated by equating  $\lambda$ -terms with the same Böhm tree, is sensible and non-extensional.  $\mathcal{B}$  is distinct from  $\mathcal{H}$  and  $\mathcal{H}^*$ , so that  $\mathcal{H} \subset \mathcal{B} \subset \mathcal{H}^*$ . Notice that not all

$\lambda$ -theories  $T$  satisfying the condition  $\mathcal{B} \subset T \subset \mathcal{H}^*$  are extensional (see the remark after Thm. 45).

In the remaining part of this section we characterize the  $\lambda$ -theory  $\mathcal{H}^*$  in terms of Böhms trees.

For all  $\lambda$ -terms  $M$  and  $N$ , we write  $M \leq_\eta N$  if  $BT(N)$  is a (possibly infinite)  $\eta$ -expansion of  $BT(M)$  (see [4, Def. 10.2.10]). For example, let  $J \equiv \Theta(\lambda jxy.x(jy))$ , where  $\Theta$  is the Turing's fixpoint combinator. Then,  $x \leq_\eta Jx$  (see [4, Example 10.2.9]), since

$$\begin{aligned} Jx &=_{\lambda\beta} \lambda z_0.x(Jz_0) =_{\lambda\beta} \lambda z_0.x(\lambda z_1.z_0(Jz_1)) \\ &=_{\lambda\beta} \lambda z_0.x(\lambda z_1.z_0(\lambda z_2.z_1(Jz_2))) =_{\lambda\beta} \dots \end{aligned}$$

The following is the Böhms tree of  $Jx$ :

$$\begin{array}{c} BT(Jx) = \lambda z_0.x \\ | \\ \lambda z_1.z_0 \\ | \\ \lambda z_2.z_1 \\ | \\ \dots \end{array}$$

We write  $N =_\eta M$  if there exists a Böhms-like tree  $A$  such that  $BT(M) \leq_\eta A$  and  $BT(N) \leq_\eta A$  (see [4, Def. 10.2.25] and the proof of the point (i  $\Rightarrow$  ii) in [4, Thm. 10.2.31]). It is well known that

$$M =_{\mathcal{H}^*} N \Leftrightarrow M =_\eta N \quad (\text{see [4, Thm. 19.2.9]}).$$

## 2.3 Graph models

The class of graph models belongs to Scott's continuous semantics. Historically, the first graph model was Scott's  $P_\omega$ , which is also known in the literature as ‘the graph model’. ‘Graph’ referred to the fact that the continuous functions were encoded in the model via (a sufficient fragment of) their graph.

As a matter of notation, for every set  $D$ ,  $D^*$  is the set of all finite subsets of  $D$ , while  $\mathcal{P}(D)$  is the powerset of  $D$ . If  $C$  is a complete partial ordering (cpo, for short), then  $[C \rightarrow C]$  denotes the cpo of all Scott continuous functions from  $C$  into  $C$ .

**Definition 1** A graph model  $D$  is a pair  $(|D|, c_D)$ , where  $|D|$  is an infinite set, called the web of  $D$ , and  $c_D : |D|^* \times |D| \rightarrow |D|$  is an injective total function.

When there is no danger of confusion, we use the same notation  $D$  for the graph model and its web. Thus, for example,  $\alpha \in D$  means  $\alpha \in |D|$ .

As a matter of notation, we write  $a \rightarrow_D \alpha$ , or also simply  $a \rightarrow \alpha$ , for  $c_D(a, \alpha)$ . When parenthesis are omitted, then association to the right is assumed. For example,  $a \rightarrow b \rightarrow \alpha$

stands for  $c_D(a, c_D(b, \alpha))$ . If  $\bar{a} = a_1 \dots a_n$  is a sequence of finite subsets of  $D$ , then we write  $\bar{a}_n \rightarrow \alpha$  for  $a_1 \rightarrow a_2 \rightarrow \dots \rightarrow a_n \rightarrow \alpha$ .

The function  $c_D$  is useful to encode a fragment of the graph of a Scott continuous function  $f : \mathcal{P}(D) \rightarrow \mathcal{P}(D)$  as a subset  $G(f)$  of  $D$ :

$$G(f) = \{a \rightarrow_D \alpha \mid \alpha \in f(a) \text{ and } a \in D^*\}. \quad (2)$$

Any graph model  $D$  is used to define a model of lambda calculus through the reflexive cpo  $(\mathcal{P}(D), \subseteq)$  determined by two Scott continuous mappings  $G : [\mathcal{P}(D) \rightarrow \mathcal{P}(D)] \rightarrow \mathcal{P}(D)$  and  $F : \mathcal{P}(D) \rightarrow [\mathcal{P}(D) \rightarrow \mathcal{P}(D)]$ . The function  $G$  is defined in (2), while  $F$  is defined as follows:

$$F(X)(Y) = \{\alpha \in D : (\exists a \subseteq Y) a \rightarrow_D \alpha \in X\}.$$

For more details we refer the reader to Berline [7] and to Chapter 5 of Barendregt's book [4].

Let  $Env_D$  be the set of  $D$ -environments  $\rho$  mapping the set of the variables of lambda calculus into  $\mathcal{P}(D)$ . If  $Y \subseteq D$ , then the environment  $\rho[x := Y]$  is defined by:  $\rho[x := Y](x) = Y$ ;  $\rho[x := Y](z) = \rho(z)$  for  $z \neq x$ . The interpretation  $M^D$  of a  $\lambda$ -term  $M$  in an environment  $\rho$  is defined as follows.

- $x_\rho^D = \rho(x)$
- $(MN)_\rho^D = \{\alpha \in D : (\exists a \subseteq N_\rho^D) a \rightarrow \alpha \in M_\rho^D\}$
- $(\lambda x.M)_\rho^D = \{a \rightarrow \alpha : \alpha \in M_{\rho[x:=a]}^D\}$

If  $\bar{x} \equiv x_1 \dots x_n$  is a sequence of variables and  $\bar{a} = a_1 \dots a_n$  is a sequence of finite subsets of  $D$ , then we have

$$(\lambda \bar{x}.M)_\rho^D = \{\bar{a}_n \rightarrow \alpha : \alpha \in M_{\rho[x_1:=a_1] \dots [x_n:=a_n]}^D\}.$$

We turn now to the interpretation of  $\Omega$  in graph models. The following remark gives a necessary condition and a sufficient condition for  $\alpha$  to be in the interpretation of  $\Omega$  in a graph model.

**Lemma 2** (Baeten-Boerboom [3]) *Let  $D$  be a graph model and  $\alpha \in D$ . Then we have:*

- (i) *If  $\alpha \in \Omega^D$ , then there exists  $a$  such that  $a \rightarrow \alpha \in a$ .*
- (ii) *If there exists  $\beta \in D$  such that  $\{\beta\} \rightarrow \alpha = \beta$ , then  $\alpha \in \Omega^D$ .*

Given a graph model  $D$ , we have that  $M^D = N^D$  if, and only if,  $M_\rho^D = N_\rho^D$  for all environments  $\rho$ . The  $\lambda$ -theory  $Th(D)$  induced by  $D$  is defined as

$$Th(D) = \{M = N : M^D = N^D\}.$$

A  $\lambda$ -theory induced by a graph model will be called a *graph theory*. The graph model  $D$  is called *sensible* if  $Th(D)$  is a sensible  $\lambda$ -theory. Kerth has shown in [23] that there exists a continuum of different (sensible) graph theories. It is well known that the graph theory  $Th(D)$  is never extensional because  $(\lambda x.x)^D \neq (\lambda xy.xy)^D$ .

Di Gianantonio and Honsell [18] have shown that graph models are related to filter models (see Coppo-Dezani [16] and Barendregt et al. [5]), since the class of graph theories is included within the class of  $\lambda$ -theories induced by non-extensional filter models. Alessi et al. [2] have shown that this inclusion is strict, namely there exists an equation between  $\lambda$ -terms, which is omitted in graph semantics, whilst it is satisfied in some non-extensional filter model.

A graph theory  $T$  will be called

1. *the minimal graph theory* if  $T \subseteq Th(D)$  for all graph models  $D$ ;
2. *the minimal sensible graph theory* if  $T$  is sensible and  $T \subseteq Th(D)$  for all sensible graph models  $D$ ;
3. *the maximal sensible graph theory* if  $T$  is sensible and  $Th(D) \subseteq T$  for all sensible graph models  $D$ .

A class  $\mathcal{C}$  of graph models *omits* (*forces*, respectively) an equation if it fails (holds) in all models of  $\mathcal{C}$ . If  $\mathcal{C}$  omits an equation  $M = N$ , then it omits all  $\lambda$ -theories including  $M = N$ .

The completion method for building graph models from ‘partial pairs’ was initiated by Longo in [25] and developed on a wide scale by Kerth in [23, 24]. This method is useful to build models satisfying prescribed constraints, such as domain equations and inequations, and it is particularly convenient for dealing with the equational theories of graph models.

**Definition 3** A partial pair  $A$  is given by an infinite set  $|A|$  and by a partial, injective function  $c_A : |A|^* \times |A| \rightarrow |A|$ .

As for graph models, we use the same notation  $A$  for the partial pair and its underlying set.

A partial pair is a graph model if and only if  $c_A$  is total. We always suppose that no element of  $A$  is a pair. This is not restrictive because partial pairs can be considered up to isomorphism.

Lambda terms can be interpreted by induction in partial pairs  $A$  ways in the obvious way. For example, we have that  $(MN)_\rho^A = \{ \alpha \in A : (\exists a \subseteq N_\rho^A) [(a, \alpha) \in \text{dom}(c_A) \wedge c_A(a, \alpha) \in M_\rho^A] \}$  and  $(\lambda x.M)_\rho^A = \{ c_A(a, \gamma) \in A : (a, \gamma) \in \text{dom}(c_A) \wedge \gamma \in M_{\rho[x:=a]}^A \}$ .

**Definition 4** Let  $A$  be a partial pair. The canonical completion of  $A$  is the graph model  $E$  defined as follows:

- $E = \bigcup_{n \in \omega} E_n$ , where  $E_0 = A$ ,  $E_{n+1} = E_n \cup ((E_n^* \times E_n) - \text{dom}(c_A))$ .



- Given  $a \in E^*$ ,  $\alpha \in E$ ,

$$c_E(a, \alpha) = \begin{cases} c_A(a, \alpha) & \text{if } c_A(a, \alpha) \text{ is defined} \\ (a, \alpha) & \text{otherwise} \end{cases}$$

It is easy to check that the canonical completion of a given partial pair  $A$  is actually a graph model. The canonical completion of a total pair  $A$  is equal to  $A$ .

A notion of *rank* can be naturally defined on the canonical completion  $E$  of a partial pair  $A$ . The elements of  $A$  are the elements of rank 0, while an element  $\alpha \in E - A$  has rank  $n$  if  $\alpha \in E_n$  and  $\alpha \notin E_{n-1}$ .

Classic graph models, such as Scott's  $P_\omega$  [4], Park's  $\mathcal{P}$  [7] and Engeler's  $\mathcal{E}_B$  (where  $B$  is an arbitrary nonempty set) [7], can be viewed as the canonical completions of suitable partial pairs. In fact,  $P_\omega$ ,  $\mathcal{P}$  and  $\mathcal{E}_B$  are respectively isomorphic to the canonical completions of  $A = (\{0\}, c_A)$  (with  $c_A(\emptyset, 0) = 0$ ),  $D = (\{p\}, c_D)$  (with  $c_D(\{p\}, p) = p$ ) and  $E = (B, c_E)$  (with  $c_E$  the empty function).

Let  $\bar{x} = x_1 \dots x_n$  be a sequence of variables and  $\rho$  be a  $D$ -environment such that  $\rho(x_i)$  is a finite set. As a matter of notation, we write  $\rho(\bar{x}_n) \rightarrow \alpha$  for  $\rho(x_1) \rightarrow \rho(x_2) \rightarrow \dots \rightarrow \rho(x_n) \rightarrow \alpha$ .

### 3 Weak product

In this section we introduce the notion of *weak product* of graph models, which is the main technical device used in the proof of the existence of the least (sensible) graph theory. The idea of a weak product is the following: given two graph models  $D_1$  and  $D_2$ , construct the partial pair whose web is the disjoint union of the webs of  $D_1$  and  $D_2$ , and whose coding function is the disjoint union of their coding functions. The canonical completion of this partial pair is the weak product of  $D_1$  and  $D_2$ .

As a matter of notations, given two sets  $A_1$  and  $A_2$ , we write  $A_1 \uplus A_2$  their disjoint union,  $in_i : A_i \rightarrow A_1 \uplus A_2$  the canonical injections and  $pr_i : 2^{A_1 \uplus A_2} \rightarrow 2^{A_i}$  the canonical projections.

**Definition 5** Let  $D_1$  and  $D_2$  be graph models. We define the partial pair  $D_1 \uplus D_2$  by

$$\begin{aligned} |D_1 \uplus D_2| &= |D_1| \uplus |D_2| \\ c_{D_1 \uplus D_2}(b, \beta) &= \begin{cases} in_i(c_{D_i}(a, \alpha)) & \text{if } b = \{in_i(\alpha') \mid \alpha' \in a\}, \beta = in_i(\alpha) \\ \text{undefined} & \text{otherwise} \end{cases} \end{aligned}$$

**Definition 6** Let  $D_1$  and  $D_2$  be graph models. The graph model  $D_1 \diamond D_2$ , called the weak product of  $D_1$  and  $D_2$ , is the canonical completion of the partial pair  $D_1 \uplus D_2$  defined above.

These definitions extend to countable products by considering countable disjoint unions of webs. Countable weak products are denoted by  $\diamond_{i \in \omega} D_i$ .

For the sake of visibility of statements and proofs, we will suppose that, when forming weak products, the factors' webs are disjoint, and that the canonical injections are replaced by set inclusions. So, for instance, if  $M$  is a  $\lambda$ -term and  $D_i$  is a factor of a weak product  $E$ , it makes sense to write  $M^{D_i} \subseteq M^E$ .

The rest of this section is devoted to the proof of the main properties of this construction:

- (i) The theory of a weak product is included in the intersection of the theories of its factors (see Section 3.1).
- (ii) The theory of a weak product is semisensible (see Section 3.2).
- (iii) The inclusion in (i) is strict in general (see Section 3.3).

### 3.1 The theory of a weak product and of its factors

In this section we show that the theory  $Th(E)$  of a weak product  $E$  is included in the theory  $Th(D_i)$  of each of its factor  $D_i$ . The idea is to prove that, for all closed  $\lambda$ -terms  $M$

$$M^{D_i} = M^E \cap D_i. \quad (3)$$

This takes a structural induction on  $M$ , and hence the analysis of open terms too. Roughly, we are going to show that equation (3) holds for open terms as well, provided that the environments satisfy a suitable closure property introduced below.

In the rest of this section,  $D_i$  is a factor of a (finite or countable) weak product  $E$ .

**Definition 7** We call *i*-flattening the function  $f_i : E \rightarrow E$  defined by induction on the rank of elements of  $E$  as follows:

if  $rank(x) = 0$  then  $f_i(x) = x$

if  $rank(x) > 0$  and  $x = (a, y)$  then

$$f_i(x) = \begin{cases} c_{D_i}(f_i(a) \cap D_i, f_i(y)) & \text{if } f_i(y) \in D_i \\ x & \text{otherwise,} \end{cases}$$

where  $f_i(a) = \{f_i(y) : y \in a\}$ .

The following easy facts will be useful:

**Fact 8** (a) For all  $x \in E$ , if  $f_i(x) \notin D_i$  then  $f_i(x) = x$ .

(b) If  $a \cup \{z\} \subseteq E$  and  $f_i(z) \in D_i$ , then  $f_i(c_E(a, z)) \in D_i$ .

We notice that Fact 8(b) holds, a fortiori, if  $z \in D_i$ .

**Definition 9** For  $a \subseteq E$  let  $\hat{a} = a \cup f_i(a)$ ; we say that  $a$  is  $i$ -closed if  $\hat{a} = a$ .

In other words,  $a$  is  $i$ -closed if  $f_i(a) \subseteq a$ .

**Lemma 10** For all  $a \subseteq E$ ,  $\hat{a} \cap D_i = f_i(a) \cap D_i$ .

**Proof.** By definition,  $\hat{a} = a \cup f_i(a)$ , hence

$$\hat{a} \cap D_i = (a \cap D_i) \cup (f_i(a) \cap D_i).$$

Since  $f_i$  restricted to  $D_i$  is the identity function, we have  $a \cap D_i \subseteq f_i(a) \cap D_i$ , and we are done. ■

**Definition 11** Let  $\rho : \text{Var} \rightarrow \mathcal{P}(E)$  be a  $E$ -environment. We define the  $i$ -restriction  $\rho_i$  of  $\rho$  by  $\rho_i(x) = \rho(x) \cap D_i$ , while we say that  $\rho$  is  $i$ -closed if for every variable  $x$ ,  $\rho(x)$  is  $i$ -closed.

The following proposition is the key technical lemma of the section:

**Proposition 12** Let  $M$  be a  $\lambda$ -term and  $\rho$  be an  $i$ -closed  $E$ -environment; then

- (a)  $M_\rho^E$  is  $i$ -closed.
- (b)  $M_\rho^E \cap D_i \subseteq M_{\rho_i}^E$ .

**Proof.** We prove (a) and (b) simultaneously by induction on the structure of  $M$ . If  $M \equiv x$ , both statements are trivially true.

Let  $M \equiv \lambda x.N$ , and let us start by proving the statement (a): given  $y \in M_\rho^E$ , we have to show that  $f_i(y) \in M_\rho^E$ . First we remark that, if  $\text{rank}(y) = 0$  or if  $y = (a, z)$  and  $f_i(z) \notin D_i$ , then by Fact 8(a)  $f_i(y) = y$  and we are done. Then, let  $y = (a, z)$  and  $f_i(z) \in D_i$ ; we have

$$\begin{array}{ll}
y \in M_\rho^E & \\
\Rightarrow z \in N_{\rho[x:=a]}^E & \text{by definition of } (\_)^E \\
\Rightarrow z \in N_{\rho[x:=\hat{a}]}^E & \text{by monotonicity of } (\_)^E \text{ w.r.t. environments} \\
\Rightarrow f_i(z) \in N_{\rho[x:=\hat{a}]}^E & \text{by (a), remark that } \rho[x:=\hat{a}] \text{ is closed} \\
\Rightarrow f_i(z) \in N_{(\rho[x:=\hat{a}])_i}^E & \text{by (b), since } f_i(z) \in D_i \\
\Rightarrow f_i(z) \in N_{\rho_i[x:=f_i(a) \cap D_i]}^E & \text{by Lemma 10} \\
\Rightarrow c_E(f_i(a) \cap D_i, f_i(z)) \in M_{\rho_i}^E & \text{by definition of } (\_)^E \\
\Rightarrow c_{D_i}(f_i(a) \cap D_i, f_i(z)) \in M_{\rho_i}^E & \text{by definition of } (E, i) \\
\Rightarrow f_i(y) \in M_{\rho_i}^E & \text{by definition of } f_i \\
\Rightarrow f_i(y) \in M_\rho^E & \text{by monotonicity of } (\_)^E
\end{array}$$

Let us prove that  $M \equiv \lambda x.N$  satisfies (b):

$$\begin{array}{ll}
y \in M_\rho^E \cap D_i & \\
\Rightarrow (\exists a \subseteq D_i)(\exists z \in D_i) y = c_{D_i}(a, z) \text{ and } z \in N_{\rho[x:=a]}^E & \text{by definition of } (-)^E \text{ and since } y \in D_i \\
\Rightarrow z \in N_{(\rho[x:=a])_i}^E & \text{by (b), remark that } \hat{a} = a \\
\Rightarrow z \in N_{\rho_i[x:=a]}^E & \text{since } a \subseteq D_i \\
\Rightarrow y \in M_{\rho_i}^E & \text{by definition of } (-)^E
\end{array}$$

Let  $M \equiv PQ$ .

(a) Let  $z \in (PQ)_\rho^E$ . If  $f_i(z) = z$  we are done, otherwise by Lemma 8(a)  $f_i(z) \in D_i$ . Moreover,  $\exists a \subseteq E$  such that  $c_E(a, z) \in P_\rho^E$  and  $a \subseteq Q_\rho^E$ . Applying (a) and Fact 8(b) we get

$$f_i(c_E(a, z)) = c_{D_i}(f_i(a) \cap D_i, f_i(z)) = c_E(f_i(a) \cap D_i, f_i(z)) \in P_\rho^E.$$

Applying (a) to  $Q$  we get  $f_i(a) \subseteq Q_\rho^E$ . Hence  $f_i(z) \in M_\rho^E$ .

(b) If  $z \in (PQ)_\rho^E \cap D_i$ , then  $\exists a \subseteq E$  such that  $c_E(a, z) \in P_\rho^E$  and  $a \subseteq Q_\rho^E$ . Since  $\rho$  is  $i$ -closed and  $z \in D_i$ , then by (a) and by Fact 8(b) we get  $f_i(c_E(a, z)) = c_{D_i}(f_i(a) \cap D_i, z) \in P_\rho^E$  and  $f_i(a) \cap D_i \subseteq Q_\rho^E$ . Now, by (b), we obtain  $c_{D_i}(f_i(a) \cap D_i, z) \in P_{\rho_i}^E$  and  $f_i(a) \cap D_i \subseteq Q_{\rho_i}^E$ , and we conclude  $z \in (PQ)_{\rho_i}^E$ . ■

**Proposition 13** Let  $M$  be a  $\lambda$ -term and  $\rho : \text{Var} \rightarrow \mathcal{P}(D_i)$  be a  $D_i$ -environment; then we have  $M_\rho^E \cap D_i = M_{\rho_i}^{D_i}$ .

**Proof.** We prove by induction on the structure of  $M$  that  $M_\rho^E \cap D_i \subseteq M_{\rho_i}^{D_i}$ . The converse is ensured by  $M_\rho^{D_i} \subseteq M_\rho^E$  and  $M_\rho^{D_i} \subseteq D_i$ , both trivially true.

If  $M \equiv x$ , the statement trivially holds.

Let  $M \equiv \lambda x.N$ ; if  $y \in M_\rho^E \cap D_i$ , then  $y = c_{D_i}(a, z)$  with  $a \cup \{z\} \subseteq D_i$ , and  $z \in N_{\rho[x:=a]}^E$ . By induction hypothesis  $z \in N_{\rho[x:=a]}^{D_i}$ , and hence  $c_{D_i}(a, z) = y \in M_{\rho_i}^{D_i}$ .

Let  $M \equiv PQ$ ; If  $z \in (PQ)_\rho^E \cap D_i$ , then  $\exists a \subseteq E$  such that  $c_E(a, z) \in P_\rho^E$  and  $a \subseteq Q_\rho^E$ . Since  $\rho$  is  $i$ -closed and  $z \in D_i$ , we can use Lemma 8(b) and Prop. 12(i) to obtain

$$f_i(c_E(a, z)) = c_{D_i}(f_i(a) \cap D_i, z) \in P_\rho^E.$$

Hence we can use the induction hypothesis to get  $c_{D_i}(f_i(a) \cap D_i, z) \in P_{\rho_i}^{D_i}$ . Moreover,  $f_i(a) \cap D_i \subseteq Q_{\rho_i}^{D_i}$  by using again Prop. 12(i) and the induction hypothesis on  $Q$ . Hence  $z \in (PQ)_{\rho_i}^{D_i}$ . ■

**Theorem 14**  $Th(E) \subseteq Th(D_i)$ .

**Proof.** Let  $M^E = N^E$ . By the previous proposition we have

$$M^{D_i} = M^E \cap D_i = N^E \cap D_i = N^{D_i}.$$

■

The existence of the least (resp. the least sensible) graph theory will be a consequence of Thm. 14 (see Section 4).

The following easy properties of weak products will be used in Section 4.2:

**Proposition 15** *Let  $E = \diamond_{i \in I} D_i$ . For all  $x \in E$  there exists a unique  $j \in I$  such that  $f_j(x) \in D_j$ .*

**Proof.** By induction on the rank of  $x$ . ■

**Proposition 16** *Let  $E = \diamond_{i \in I} D_i$  and  $M$  be a closed  $\lambda$ -term. For all  $x \in M^E$  there exists a unique  $j \in I$  such that  $f_j(x) \in M^{D_j}$ .*

**Proof.** By Prop. 15 we know that there is a unique  $j$  such that  $f_j(x) \in D_j$ , while By Prop. 12(a) we have that  $f_j(x) \in M^E$ . The conclusion follows from Prop. 13. ■

### 3.2 The theory of a weak product is semisensible

In this section we show that *stratified* graph models have semisensible theories. A graph model is stratified if it is the completion of a proper partial pair, i.e. one whose coding function is not total. Since weak products are particular stratified graph models, then the theory of a weak product is also semisensible.

Semisensibility of the theory of a stratified graph model is proved by case analysis, on the order of unsolvable terms (see Def. 23 for the definition of order of an unsolvable). The fact that unsolvables of order 0 cannot be equated to a solvable in a stratified graph model is shown in Lemma 24 by using the approximation theorem below.

Concerning unsolvable of finite order, we introduce the notion of *height* of elements of the model, and then rely on the previous case (Lemma 26).

For the unsolvable of infinite order, we rely on a general property of graph models, their non-extensionality, to show that such terms cannot be equated to solvables in any graph model (Lemma 28).

#### 3.2.1 An Approximation Theorem

Approximation theorems are an important tool in the analysis of the  $\lambda$ -theories induced by models of lambda calculus. In this section we provide an approximation theorem for the class of stratified graph models: we show that the interpretation of a  $\lambda$ -term in a stratified graph model is the union of the interpretations of its direct approximants. This approximation theorem will be applied in Section 3.2.2 to show that the interpretation of an unsolvable of order 0 in a stratified graph model is a set of elements of rank 0. We do not claim any particular

originality for the approximation theorem we prove in this section, since it is a very similar to that in [21] and it is a particular case of that in [6]. However, for the sake of completeness, we provide a proof.

Let  $D$  be a stratified graph model, which is the completion of the partial pair  $A$ . Recall that  $D_0 = A$  and  $D_{n+1} = (D_n^* \times D_n) - \text{dom } c_A$ . For every  $X \subseteq D$ , we denote by  $X_{\underline{n}} = X \cap D_n$ .

The underlined natural numbers  $\underline{n}$  are called *labels*. Lambda terms with occurrences of labels are called *labelled-terms*. For example,  $(\lambda x.x_{\underline{n}})_{\underline{m}}y$  and  $(y_{\underline{n}})_{\underline{m}}$  are labelled-terms. Note that the set of ordinary  $\lambda$ -terms is a proper subset of the set of labelled terms (those without any label). If  $N$  is a labelled term, we denote by  $|N|$  the  $\lambda$ -term obtained by erasing all labels of  $N$ . For example, we have that  $|(\lambda x.x_{\underline{n}})_{\underline{m}}y| = (\lambda x.x)y$ .

Labelled terms are interpreted in  $D$ : the interpretation function of labelled terms is the unique extension of the interpretation function of  $\lambda$ -terms such that, for every labelled term  $M$  and label  $\underline{n}$ ,  $(M_{\underline{n}})_\rho^D = (M_\rho^D)_{\underline{n}}$ .

As a matter of notation, we write  $M =_{D,\rho} N$  for  $M_\rho^D = N_\rho^D$  and  $M \subseteq_{D,\rho} N$  for  $M_\rho^D \subseteq N_\rho^D$ .

An easy fact that we will use later is that, for all labelled terms  $M, N$  and environment  $\rho$ , if  $N$  is obtained by erasing some of the labels of  $M$ , then,  $M \subseteq_{D,\rho} N$ . In particular, for every labelled term  $M$  and environment  $\rho$ ,  $M \subseteq_{D,\rho} |M|$ .

**Definition 17** *The weak direct approximant (w.a.) of a  $\lambda$ -term is defined by induction as follows:*

- $x^{wa} = x$ ;
- $(\lambda x.M)^{wa} = \lambda x.M^{wa}$ ;
- $(MN)^{wa} = M^{wa}N^{wa}$  if  $MN$  is not a redex;
- $((\lambda x.M)N)^{wa} = (\lambda x.M^{wa})_{\underline{0}}N^{wa}$ .

The weak direct approximant  $M^{wa}$  of a  $\lambda$ -term  $M$  is a labelled term such that  $|M^{wa}| = M$ . Moreover, it is easy to show that  $M^{wa} \subseteq_{D,\rho} M$  for every  $\lambda$ -term  $M$  and environment  $\rho$ , so that we have

$$\bigcup \{(N^{wa})_\rho^D : M =_{\lambda\beta} N\} \subseteq M_\rho^D.$$

The remaining part of this section is devoted to prove that the inclusion above is actually an equality.

**Theorem 18** (The Approximation Theorem) *Let  $D$  be a stratified graph model. For every  $\lambda$ -term  $M$  and environment  $\rho$ , we have*

$$M_\rho^D = \bigcup \{(N^{wa})_\rho^D : M =_{\lambda\beta} N\}.$$

**Proof.** The proof is divided into claims.

We say that a labelled-term  $N$  is *completely labelled* if every subterm of  $N$  has at least a label. For example,  $((\lambda x.x_{\underline{n}})_{\underline{0}} y_{\underline{m}})_{\underline{0}}$  and  $((\lambda x.x_{\underline{n}})_{\underline{0}} (y_{\underline{m}})_{\underline{0}})_{\underline{0}}$  are two completely labelled versions of the  $\lambda$ -term  $(\lambda x.x)y$ .

**Claim 19** *For every  $\lambda$ -term  $M$  and for every environment  $\rho$  we have:*

$$M_{\rho}^D = \bigcup \{N_{\rho}^D : N \text{ is a completely labelled term, } |N| = M\}.$$

It is sufficient to show by induction on  $M$  that, if  $\alpha \in (M)_{\rho}^D \cap D_n$ , then there is a completely labelled term  $N$  such that  $|N| = M$  and  $\alpha \in (N)_{\rho}^D$ .

**Claim 20** *The rewriting system generated by the rules*

$$(\lambda x.P)_{\underline{n+1}} Q \rightarrow_{lab} P_{\underline{n}}[x := Q_{\underline{n}}]; \quad (P_{\underline{n}})_{\underline{m}} \rightarrow_{lab} P_{\min(n,m)}$$

*is Church-Rosser and strongly normalizing.*

The proof is in Section 14.1 of Barendregt's book [4]; remark that:

- if  $M, N$  are labelled terms and  $M \rightarrow_{lab}^* N$ , then  $|M| \rightarrow_{\beta}^* |N|$ .
- every  $\rightarrow_{lab}$  reduct of a completely labelled term is completely labelled.
- the usual substitution lemma holds for labelled terms: for all labelled terms  $P$  and  $Q$  and environment  $\rho$ ,  $(P[x := Q])_{\rho}^D = P_{\rho[x := Q_{\rho}^D]}^D$ .

The next claim shows that the interpretation of a labelled term does not decrease along  $\rightarrow_{lab}$  reduction paths:

**Claim 21** *For all labelled  $\lambda$ -terms  $P$  and  $Q$  and environment  $\rho$ ,*

$$(\lambda x.P)_{\underline{n+1}} Q \subseteq_{D, \rho} P_{\underline{n}}[x := Q_{\underline{n}}]$$

Let  $\alpha \in ((\lambda x.P)_{\underline{n+1}} Q)_{\rho}^D$ . Then there exist  $b \subseteq D$  and  $\alpha \in D$  such that  $b \rightarrow \alpha \in ((\lambda x.P)_{\underline{n+1}})_{\rho}^D$  and  $b \subseteq Q_{\rho}^D$ . Hence  $b \cup \{\alpha\} \subseteq |D|_n$ , and  $\alpha \in P_{\rho[x := b]}^D$ . By these two last relations and by  $b \subseteq (Q_{\underline{n}})_{\rho}^D$  we obtain that  $\alpha \in (P_{\underline{n}})_{\rho[x := (Q_{\underline{n}})_{\rho}^D]}^D$ . By the substitution lemma we conclude that  $\alpha \in P_{\underline{n}}[x := Q_{\underline{n}}]_{\rho}^D$ .

Finally, the approximation theorem:

**Claim 22** *For all  $\lambda$ -terms  $M$  and environment  $\rho$ ,*

$$M_{\rho}^D = \bigcup \{(N^{wa})_{\rho}^D : M =_{\lambda\beta} N\}.$$

Let  $N$  be a completely labelled term such that  $|N| = M$ . By Claim 21 we get  $N \subseteq_{D,\rho} N_1$ , where  $N_1$  is the normal form of  $N$  w.r.t. the rewriting rules  $\rightarrow_{lab}$ . Since  $N_1$  has no redexes w.r.t.  $\rightarrow_{lab}$ , and it is completely labelled as remarked above, then every redex of the  $\lambda$ -term  $|N_1|$  should occur in  $N_1$  as  $(\lambda x.P)_{\underline{0}}Q$ . Let  $N_2$  be the  $\lambda$ -term obtained from  $N_1$  by erasing all labels  $n > 0$ ; we have  $N_1 \subseteq_{D,\rho} N_2$ . Finally, we get a new term  $N_3$  by erasing from  $N_2$  all occurrences of the label  $\underline{0}$  which are not in the position  $(\lambda x.P)_{\underline{0}}Q$ . Note that  $N_3$  is the direct approximant of  $|N_3|$ .

In conclusion, we have

$$N \subseteq_{D,\rho} N_1 \subseteq_{D,\rho} N_2 \subseteq_{D,\rho} N_3; \quad N_3 = |N_3|^{wa}. \quad (4)$$

Moreover, as remarked above, we also have

$$M \rightarrow_{\beta}^* |N_3|. \quad (5)$$

In conclusion,

$$\begin{array}{ll} M = \bigcup \{N_{\rho}^D : N \text{ completely labelled, } |N| = M\} & \text{by Claim 19} \\ \subseteq_{D,\rho} \bigcup \{Q^{wa} : M =_{\beta} Q\} & \text{by (4) and (5)} \\ \subseteq_{D,\rho} M & \text{as remarked after Definition 17.} \end{array}$$

This concludes the proof of the approximation theorem. ■

### 3.2.2 The theory of a stratified graph model

We apply the approximation theorem to show that stratified graph models have semisensible theories. Let us recall the definition of *order* of an unsolvable  $\lambda$ -term:

**Definition 23** *An unsolvable  $\lambda$ -term  $U$  has*

1. *order 0 if it is not  $\beta$ -equivalent to an abstraction term;*
2. *order  $n$  if  $U =_{\lambda\beta} \lambda x_1 \dots x_n.T$  and  $T$  has order 0;*
3. *order  $\omega$  if it has no finite order.*

For example,  $\Omega$  and  $\Omega_3$  are unsolvable of order 0,  $\lambda x.\Omega$  has order 1, while  $Y\mathbf{k}$  has order  $\omega$ , where  $Y$  is any fixpoint combinator.

**Lemma 24** *Let  $D$  be a stratified model, and  $U$  be an unsolvable of order 0. Then, for every environment  $\rho$ , we have:*

$$U_{\rho}^D \subseteq D_0.$$



**Proof.** If  $N =_{\lambda\beta} U$  then  $N$  is also an unsolvable of order 0. Hence,  $N \equiv (\lambda x.P)Q_1 \dots Q_m$ , so that  $N^{wa} \equiv (\lambda x.P^{wa})_{\underline{0}}Q_1^{wa} \dots Q_m^{wa}$ . The conclusion follows from the approximation theorem because  $((\lambda x.P^{wa})_{\underline{0}}Q_1^{wa} \dots Q_m^{wa})_{\rho}^D \subseteq D_0$ . ■

An easy corollary of this lemma is that, in stratified graph models, unsolvables of order 0 cannot be equated to solvables, since the interpretation of any solvable contains elements of arbitrary rank (see Lemma 27).

In order to deal with unsolvable of arbitrary order, we introduce the notion of *height* in a stratified model.

**Definition 25** Let  $D$  be a stratified model and  $\alpha \in D$ . Then we define by induction over the rank the notion of height  $h(\alpha)$  of  $\alpha$ :

- If  $\text{rank}(\alpha) = 0$ , then  $h(\alpha) = 0$ ;
- If  $\text{rank}(\alpha) > 0$  and  $\alpha = (b, \beta)$ , then  $h(\alpha) = 1 + h(\beta)$ .

Notice that, whenever  $\alpha = a_1 \rightarrow \dots \rightarrow a_n \rightarrow \beta$  with  $\text{rank}(\beta) = 0$ , then  $h(\alpha) \leq n$ .

**Lemma 26** Let  $D$  be a stratified model and  $U$  be an unsolvable of order  $n$ . Then, for every environment  $\rho$ , we have:

$$\alpha \in U_{\rho}^D \Rightarrow h(\alpha) \leq n.$$

**Proof.** By hypothesis  $U =_{\lambda\beta} \lambda x_1 \dots x_n.T$  with  $T$  of order 0. If  $\alpha \in U_{\rho}^D$  then  $\alpha = a_1 \rightarrow \dots \rightarrow a_n \rightarrow \beta$  and  $\beta \in T_{\sigma}^D$ , where  $\sigma(x_i) = a_i$  and  $\sigma(y) = \rho(y)$  for all  $y \neq x_i$ . By Lemma 24 we have that  $\text{rank}(\beta) = 0$ . Then the conclusion follows by the remark after Def. 25. ■

Lemmata 24 and 26 show that, for any unsolvable  $U$  of finite order, the interpretation of  $U$  in a stratified graph model contains only elements whose height is not bigger than the order of  $U$ .

The next lemma shows that the interpretation of any solvable terms contains element of arbitrary height:

**Lemma 27** Let  $D$  be a stratified model and  $S \in \Lambda^o$  be a solvable  $\lambda$ -term. Then, for every natural number  $k$ , there is  $\alpha \in D$  such that  $\alpha \in S^D$  and  $h(\alpha) \geq k$ .

**Proof.** Let  $S =_{\lambda\beta} \lambda x_1 \dots x_n.x_j.P_1 \dots P_m$ . It is easy to show that  $\alpha = \emptyset^{j-1} \rightarrow (\emptyset^m \rightarrow \beta) \rightarrow \emptyset^{n-j} \rightarrow \beta \in S^D$  for all  $\beta \in D$ . If we choose  $h(\beta) = k$ , then  $h(\alpha) \geq h(\beta) = k$  and we get the conclusion. ■

So far, we have seen that in a stratified model the interpretation of an unsolvable term of finite order is different from the interpretation of any solvable term.

We show now that unsolvable terms of infinite order cannot be consistently equated to solvable terms in graph models.

**Lemma 28** *Let  $D$  be a graph model,  $U \in \Lambda^\circ$  be an unsolvable  $\lambda$ -term of infinite order and  $S \in \Lambda^\circ$  be a solvable  $\lambda$ -term. Then*

$$U^D \neq S^D.$$

**Proof.** Assume, by the way of contradiction, that  $U^D = S^D$ . Since  $S$  is solvable, there exist  $\lambda$ -terms  $M_1, \dots, M_k$  such that  $SM_1 \dots M_k =_{\lambda\beta} x$ , for an arbitrary variable  $x$ . Then we have, for any environment  $\rho$ ,

$$x =_{D,\rho} UM_1 \dots M_k.$$

Since  $U$  is unsolvable of infinite order, then  $UM_1 \dots M_k$  is also an unsolvable of infinite order. This implies that  $UM_1 \dots M_k =_{\lambda\beta} \lambda y.T$  for suitable  $y$  and  $T$ . However, the equation  $x = \lambda y.T$  does not hold in any graph model: consider an environment  $\rho$  such that  $\rho(x) = \{a \rightarrow \alpha\}$  for given finite  $a$  and  $\alpha \in D$ . Then  $(\lambda y.T)_\rho^D = \{a \rightarrow \alpha\}$ . This is not possible because, for all finite  $b \subseteq D$ , we have that  $(b \cup a) \rightarrow \alpha \in (\lambda y.T)_\rho^D$ . Contradiction. ■

Summing up, we have proved the following result:

**Theorem 29** *The theory of any stratified graph model is semisensible.*

**Corollary 30** *The theory of any weak product is semisensible.*

### 3.3 Self weak product

Thm. 14 states that the theory of a weak product is included in the intersection of those of its factors. In this section we show that this inclusion is strict in general. Moreover, in Thm. 32 below we show that self weak products do not preserve in general equations between unsolvable terms. Then it is not in general true that  $Th(D \diamond D) = Th(D)$ , whenever  $Th(D)$  is semisensible.

**Proposition 31** *Let  $D$  be a graph model satisfying the equation  $\Omega = \mathbf{i}$ . The model  $D \diamond D$ , that we call self weak product of  $D$ , does not satisfy  $\Omega = \mathbf{i}$ .*

**Proof.** By Cor. 30 the theory of  $D \diamond D$  is semisensible. ■

**Theorem 32** *There exists a graph model  $D$  satisfying the following two conditions:*

- (i)  $D \models \Omega = \lambda x.\Omega$
- (ii)  $D \diamond D \not\models \Omega = \lambda x.\Omega$ .

**Proof.** The proof is divided into claims. For the sake of clarity, we denote by  $D_a$  the first copy of  $D$  in  $D \diamond D$  and by  $D_b$  the second copy. Moreover, we assume that these (isomorphic) copies are disjoint.

Recall that every weak product is a stratified graph model.

**Claim 33** *Let  $D$  be a graph model and let  $E \equiv D \diamond D$  be the self weak product of  $D$ . Then we have:*

$$\Omega^D = \emptyset \iff E \models \Omega = \lambda x. \Omega.$$

( $\Leftarrow$ ) Assume  $\Omega^D \neq \emptyset$ . Then by Prop. 13 we have that  $\Omega^{D_a} = \Omega^E \cap D_a$ . Thus the hypothesis implies  $\Omega^E \neq \emptyset$ . Let  $\beta \in \Omega^E$  be an arbitrary element and let  $b \subseteq E$  be a finite set containing elements of rank 1. Then  $b \rightarrow \beta \in (\lambda x. \Omega)^E = \{a \rightarrow \alpha : \alpha \in \Omega^E\}$ . In conclusion, by Lemma 24 we have that  $\Omega^E \subseteq E_0$  is a set of elements of rank 0, while  $(\lambda x. \Omega)^E$  contains elements of rank greater than 0. We get the conclusion  $E \not\models \Omega = \lambda x. \Omega$ .

( $\Rightarrow$ ) The conclusion follows from the following relations:  $\Omega^E \subseteq E_0$  (see Lemma 24);  $E_0 = D_a \cup D_b$ ;  $\emptyset = \Omega^{D_i} = \Omega^E \cap D_i$  ( $i = a, b$ ) (see Prop. 13).

This concludes the proof of Claim 34.

**Claim 34** *There exists a graph model  $D$  satisfying the following two conditions:*

1.  $D \models \Omega = \lambda x. \Omega$ ;
2.  $\Omega^D \neq \emptyset$ .

We construct a graph model by using the technique of forcing introduced by Baeten-Boerboom in [3]. In the following proof we follow [8].

Let  $D$  be any infinite countable set. We are going to define by "forcing" the injective total function  $c_D : D^* \times D \rightarrow D$

We fix an enumeration of  $D$ , and an enumeration of  $D^* \times D$ . Let  $p$  be the first element in the enumeration of  $D$ .

We are going to build an infinite sequence of elements  $\alpha_n \in D \cup \{v\}$  ( $n \geq 0$ ), where  $v$  is some new element, and an infinite sequence of partial pairs  $A_n$  ( $n \geq 1$ ) such that  $|A_n|$  is a finite set and  $c_{A_n} \subseteq c_{A_{n+1}}$  (i.e., the graph of  $c_{A_n}$  is contained within the graph of  $c_{A_{n+1}}$ ).  $D$  becomes a graph model by defining  $c_D =_{def} \bigcup_{n \in \omega} c_{A_n}$ .

We start from  $|A_1| = \{p\}$ ,  $c_{A_1}(\{p\}, p) = p$  and  $\alpha_0 = p$  (note that the canonical completion of the partial pair  $A_1$  is Park's model (see Section 2.3)). It is not difficult to verify that  $\Omega^{A_1} = \{p\} = (\lambda x. \Omega)^{A_1}$  (recall that the interpretation of a  $\lambda$ -term in a partial pair is defined in Section 2.3).

Assume that the partial pair  $A_n$  and  $\alpha_0, \dots, \alpha_{n-1}$  have been built.

Let  $\alpha_n$  be the first element of  $(\lambda x. \Omega)^{A_n} - \{\alpha_0, \dots, \alpha_{n-1}\}$  if this set is non-empty, and  $v$  otherwise.

Let  $(b_n, \delta_n)$  be the first element in  $D^* \times D - \text{dom}(c_{A_n})$  and  $\gamma_n$  be the first element in  $D - (\text{range}(c_{A_n}) \cup b_n)$ .

**Case 1.**  $\alpha_n = v$ . Then  $|A_{n+1}| = |A_n| \cup b_n \cup \{\delta_n, \gamma_n\}$  and  $c_{A_{n+1}}$  is a proper extension of  $c_{A_n}$  defined as follows in the new pair  $(b_n, \delta_n)$ :

$$c_{A_{n+1}}(b_n, \delta_n) = \gamma_n$$

**Case 2.**  $\alpha_n \in D$ . Then  $|A_{n+1}| = |A_n| \cup b_n \cup \{\delta_n, \gamma_n, \beta_n, \alpha_n\}$  and  $c_{A_{n+1}}$  is a proper extension of  $c_{A_n}$  defined as follows in the new pairs  $(b_n, \delta_n)$  and  $(\{\beta_n\}, \alpha_n)$ :

$$c_{A_{n+1}}(b_n, \delta_n) = \gamma_n; \quad c_{A_{n+1}}(\{\beta_n\}, \alpha_n) = \beta_n,$$

where  $\beta_n$  is the first element of  $D$  such that :

$$\begin{aligned} (\{\beta_n\}, \alpha_n) &\in D^* \times D - (dom(c_{A_n}) \cup \{(b_n, \delta_n)\}) \text{ and} \\ \beta_n &\in D - (range(c_{A_n}) \cup \{\gamma_n\}). \end{aligned}$$

It is clear that  $c_{A_n}$  is a strictly increasing sequence of well-defined partial injective maps and that  $c_D = \cup c_{A_n}$  is total.

There remains to see that the graph model  $D$  satisfies the equation  $\Omega = \lambda x. \Omega = B$ , where  $B =_{def} \{\alpha_n : n \in \omega\} \cap D$ .

$B \subseteq (\lambda x. \Omega)^D$  follows from  $\alpha_0 = p \in (\lambda x. \Omega)^{A_1}$ , from the definition of  $\alpha_n$  ( $n > 0$ ) and from the fact that  $(\lambda x. \Omega)^{A_n} \subseteq (\lambda x. \Omega)^D$ .

$(\lambda x. \Omega)^D \subseteq B$ : suppose  $\gamma \in (\lambda x. \Omega)^D$ ; then  $\gamma \in (\lambda x. \Omega)^{A_m}$  for some  $m$  (and for all the larger ones). If  $\gamma \notin B$  then, for all  $n \geq m$ ,  $\alpha_n \neq v$  (i.e.,  $\alpha_n \in D$ ) is smaller than  $\gamma$  in the enumeration of  $D$ , contradicting the fact that there is only a finite number of such elements.

$B \subseteq \Omega^D$  :  $\alpha_n \in \Omega^D$  follows immediately from  $\alpha_0 = p \in \Omega^{A_1} \subseteq \Omega^D$ , from the fact that  $c_D(\{\beta_n\}, \alpha_n) = \beta$  and from Lemma 2.

$\Omega^D \subseteq B$  : if  $\varepsilon \in \Omega^D$  then there is an  $a \in D^*$  such that  $c_D(a, \varepsilon) \in a$  (by Lemma 2). Since  $c_D = \cup c_{A_n}$ , then either  $\varepsilon = \gamma_n$  or  $\varepsilon = \alpha_n$  for some  $n$ . Because of the choices of the  $\gamma_n$ , the first possibility is not possible.

This concludes the proof of Claim 34.

The conclusion of the theorem is now a simple corollary of Claim 33 and Claim 34. ■

**Corollary 35** *There exist graph models  $D$  satisfying the following condition:*

$$Th(D \diamond D) \neq Th(D) \cap T, \quad \text{for every sensible } \lambda\text{-theory } T.$$

## 4 Weak product and graph theories

In this section we show the existence of a minimal graph theory and of a minimal sensible graph theory. The main technical device is that of weak product studied in the above section.

## 4.1 The minimal graph theory

Let  $I$  be the set of equations between  $\lambda$ -terms which fail to hold in some graph model. For every equation  $e \in I$ , we consider a fixed graph model  $D_e$ , where the equation  $e$  fails to hold.

Then, we consider the weak product  $E = \diamond_{e \in I} D_e$ .

By Thm. 14,  $Th(E) \subseteq Th(D_e)$ , for all  $e \in I$ . In particular,  $e \notin Th(E)$ , for all  $e \in I$ ; hence:

**Theorem 36** *The theory of the graph model  $E$  is the minimal graph theory.*

## 4.2 The minimal sensible graph theory

We proceed as before: let  $I_s$  be the set of equations which fail to hold in some sensible graph model. For every  $e \in I_s$ , let  $D_e$  be a sensible graph model where the equation  $e$  fails to hold.

Then, we consider the weak product  $E_s = \diamond_{e \in I_s} D_e$ .

By Thm. 14 the theory  $Th(E_s)$  is contained within any sensible graph theory. If  $Th(E_s)$  is sensible, then we are done.

In the remaining part of this section we show that  $Th(E_s)$  is actually sensible.

The proof of the following lemma can be found in Example 5.3.7 of Kerth's thesis [22].

**Lemma 37** (Kerth [22]) *Let  $D$  be a graph model. If  $\alpha \in (\Omega_3)^D$ , then there exists a natural number  $k \geq 1$  such that*

$$\alpha = b_1 \rightarrow \dots \rightarrow b_k \rightarrow \alpha$$

*for suitable finite subsets  $b_i$  contained in the interpretation of  $\lambda x.xxx$ .*

**Lemma 38** *If all closed unsolvable  $\lambda$ -terms have the same interpretation in a graph model, then it must be the empty set.*

**Proof.** Let  $D$  be a graph model and let  $X$  be a nonempty subset of  $D$ , that is the common interpretation of all closed unsolvables. Since  $\Omega$  and  $\lambda x.\Omega$  are both unsolvables, then we have that

$$X = (\lambda x.\Omega)^D = \{a \rightarrow \alpha : \alpha \in \Omega^D\} = \{a \rightarrow \alpha : \alpha \in X\}. \quad (6)$$

It follows that  $a \rightarrow \alpha \in X$  for all finite subsets  $a$  of  $D$  and all  $\alpha \in X$ . Let  $\gamma$  be an element of  $X$ . Then  $a \rightarrow \gamma \in (\Omega_3)^D$  by (6), since  $\Omega_3$  is unsolvable and  $(\Omega_3)^D = X$ . From Lemma 37 it follows that

$$a \rightarrow \gamma = b_1 \rightarrow \dots \rightarrow b_k \rightarrow a \rightarrow \gamma,$$

where  $b_1, \dots, b_k$  are finite subsets contained in the interpretation of  $\lambda x.xxx$ . It follows that  $b_1 = a$ . By the arbitrariness of  $a$  we can conclude that  $(\lambda x.xxx)^D = D$ . This is not possible, because, for example,  $\emptyset \rightarrow \beta \notin (\lambda x.xxx)^D$ . ■

**Theorem 39** *The theory of  $E_s$  is the minimal sensible graph theory.*

**Proof.** By construction,  $Th(E_s)$  is contained within any sensible graph theory. In order to prove that  $Th(E_s)$  is sensible, let us suppose that a closed unsolvable term  $U$  has a non-empty interpretation in  $E_s$ , i.e., there exists  $\alpha \in U^{E_s}$ . By Prop. 15 there exists a unique  $e \in I_s$  such that  $f_e(\alpha) \in D_e$ . By Prop. 12(a) we have that  $f_e(\alpha) \in U^{E_s}$ , and finally, by Prop. 12(b), that  $f_e(\alpha) \in U^{D_e}$ . Since  $D_e$  is sensible, this is impossible by Lemma 38. Hence  $U^{E_s} = \emptyset$  for any closed unsolvable  $U$  (and actually for any unsolvable in any environment). ■

## 5 The minimal graph theory is not $\lambda\beta$

A longstanding open problem is whether there exists a non-syntactic model of lambda calculus whose equational theory is equal to the least  $\lambda$ -theory  $\lambda\beta$ . In Thm. 41 below we show that this model cannot be found within graph semantics. This result negatively answers Question 1 in [7, Section 6.2] for the restricted class of graph models.

We start with a lemma.

**Lemma 40** *All graph models satisfy the inequality  $\Omega_3 \leq \lambda y.\Omega_3 y$ .*

**Proof.** Let  $D$  be an arbitrary graph model and  $\alpha \in (\Omega_3)^D$ . From Lemma 37 it follows that there exists a natural number  $k \geq 1$  such that  $\alpha = b_1 \rightarrow b_2 \rightarrow \dots \rightarrow b_k \rightarrow \alpha$  for suitable finite subsets  $b_i$  contained in the interpretation of  $\lambda x.xxx$ . We have that  $\alpha = b_1 \rightarrow b_2 \rightarrow \dots \rightarrow b_k \rightarrow \alpha \in (\lambda y.\Omega_3 y)^D$  iff there exists a finite set  $d$  such that  $d \rightarrow b_2 \rightarrow \dots \rightarrow b_k \rightarrow \alpha \in (\Omega_3)^D$  and  $d \subseteq b_1$ . This last relation is true by defining  $d \equiv b_1$ , so that  $\alpha \in (\lambda y.\Omega_3 y)^D$ . In conclusion, we get  $(\Omega_3)^D \subseteq (\lambda y.\Omega_3 y)^D$ . ■

**Theorem 41** *There exists no graph model whose equational theory is  $\lambda\beta$ .*

**Proof.** Assume that there exists a graph model  $D$  whose equational theory is  $\lambda\beta$ . By Cor. 2.4 in [30] the denotations of two non- $\lambda\beta$ -equivalent closed  $\lambda$ -terms must be incomparable in every model of lambda calculus whose equational theory is  $\lambda\beta$ . Then, for all closed  $\lambda$ -terms  $M$  and  $N$  such that  $M \not\equiv_{\lambda\beta} N$ , we have that neither  $M^D \subseteq N^D$  nor  $N^D \subseteq M^D$ . We get a contradiction because of Lemma 40. ■

In Thm. 36 we have shown that there exists a minimal graph theory. By Thm. 41 we have that  $\lambda\beta$  is strictly included within the minimal graph theory. Thus, there exist equations between non- $\lambda\beta$ -equivalent terms satisfied by all graph models. In Thm. 43, whose proof is based on technical results by Selinger [30], we characterize an equation of this kind.

Let  $f$  be any  $\lambda$ -term satisfying, via a fixpoint combinator, the recursive equation  $fxy =_{\lambda\beta} fx(fx(fxy))$  for variables  $x, y$  (in other words, any three applications of  $fx$  are equivalent

to a single application) and let  $A \equiv \lambda xyzwv.fx(fy(fz(fwv)))$ . The  $\lambda$ -terms  $f$  and  $A$  were defined by Selinger in [30]. In [30, Prop. 2.1] Selinger has shown that

$$Axxxy =_{\lambda\beta} Axyyy \quad (7)$$

and

$$Axxxy \neq_{\lambda\beta} Axyyy. \quad (8)$$

This inequality has a ingenious proof based on the notion of a finite lambda reduction model.

For the sake of completeness, we recall Lemma 2.2 in [30] that will be used in the proof of Thm. 43.

**Lemma 42** (Selinger [30]) *Let  $P_1, \dots, P_n$  be  $\lambda$ -terms that are distinct in  $\lambda\beta$ , and let  $x$  be a variable not free in  $P_1, \dots, P_n$ . Then, for all terms  $M, N$  for which  $x$  is not free in  $M$  and  $N$ , and for variables  $y_1, \dots, y_n$ , we have:*

$$M(xP_1)(xP_2) \dots (xP_n) =_{\lambda\beta} N(xP_1)(xP_2) \dots (xP_n) \Rightarrow My_1y_2 \dots y_n =_{\lambda\beta} Ny_1y_2 \dots y_n.$$

As a matter of notation, let  $t \equiv \Omega_3$  and  $u \equiv \lambda y.\Omega_3y$  in the following theorem.

**Theorem 43** *Let  $T$  be the minimal graph theory (whose existence has been shown in Thm. 36). Then we have, for a variable  $x$ ,*

$$A(xt)(xt)(xt)(xu) =_T A(xt)(xt)(xu)(xu), \quad (9)$$

while

$$A(xt)(xt)(xt)(xu) \neq_{\lambda\beta} A(xt)(xt)(xu)(xu). \quad (10)$$

**Proof.** By compatibility, by  $t \leq u$  (see Lemma 40) and by (7) we obtain that the following relations hold in every graph model:

$$A(xt)(xt)(xt)(xu) \leq A(xt)(xt)(xu)(xu) \leq A(xt)(xu)(xu)(xu) =_{\lambda\beta} A(xt)(xt)(xt)(xu).$$

It easily follows (9). It remains to show the inequality (10). Assume, by the way of contradiction, the opposite:  $A(xt)(xt)(xt)(xu) =_{\lambda\beta} A(xt)(xt)(xu)(xu)$ . We can apply the hypotheses of Lemma 42 to  $M \equiv \lambda xy.Axxxy$ ,  $N \equiv \lambda xy.Axyyy$ ,  $P_1 \equiv t$  and  $P_2 \equiv u$ . Then we get the conclusion of Lemma 42:  $My_1y_2 =_{\lambda\beta} Ny_1y_2$ , that implies  $Ay_1y_1y_1y_2 =_{\lambda\beta} Ay_1y_1y_2y_2$ . This contradicts (8). ■

## 6 Omitting equations and theories

In this section we prove the main results of the paper:

- The  $\lambda$ -theory  $\mathcal{B}$  of Böhm trees is the greatest sensible graph theory.

- Graph semantics omits all equations  $M = N$  between  $\lambda$ -terms which do not have the same Böhm tree, but have the same Böhm tree up to (possibly infinite)  $\eta$ -equivalence.

We recall that the theory  $Th(D)$  of a model of lambda calculus  $D$  is the set of all equations  $M = N$  between  $\lambda$ -terms  $M$  and  $N$  which have the same interpretation in the model. A semantics  $\mathcal{C}$  of lambda calculus is *incomplete* if there exists a  $\lambda$ -theory  $T$  such that  $T \neq Th(D)$  for all models  $D \in \mathcal{C}$ . In such a case we say that the semantics *omits* the  $\lambda$ -theory  $T$ . More generally, a semantics *omits* (*forces*, respectively) an equation if it fails (holds) in all models of the semantics. If a semantics omits an equation  $M = N$ , then it omits all  $\lambda$ -theories including  $M = N$ . It is easy to verify that the set of equations ‘forced’ by a semantics  $\mathcal{C}$  constitutes a  $\lambda$ -theory. It is the minimal  $\lambda$ -theory of  $\mathcal{C}$  if it is induced by a model of  $\mathcal{C}$ .

The following two theorems are the main results of the paper. The proof of Thm. 44 is postponed to the next section.

**Theorem 44** *The graph semantics omits all equations  $M = N$  satisfying the following conditions:*

$$M =_{\mathcal{H}^*} N \text{ and } M \neq_{\mathcal{B}} N. \quad (11)$$

In other words, graph semantics omits all equations  $M = N$  between  $\lambda$ -terms which do not have the same Böhm tree, but have the same Böhm tree up to (possibly infinite)  $\eta$ -equivalence (see Section 2.2 in this paper and Barendregt [4, Section 10]).

**Theorem 45** *The  $\lambda$ -theory  $\mathcal{B}$  is the unique maximal sensible graph theory.*

**Proof.**  $\mathcal{B}$  is the equational theory of Scott’s graph model  $P_\omega$  (see Section 19.1 in [4]) and of Engeler’s graph model  $\mathcal{E}_A$  (see [7]). Let  $T$  be a sensible graph theory and suppose  $M =_T N$ . We have that  $M =_{\mathcal{H}^*} N$ , because  $\mathcal{H}^*$  is the unique maximal sensible  $\lambda$ -theory. Since graph semantics does not omit the equation  $M = N$ , then from  $M =_{\mathcal{H}^*} N$  and from Thm. 44 it follows that  $M =_{\mathcal{B}} N$ , so that  $T \subseteq \mathcal{B}$ . ■

It is well known that every graph theory is non-extensional (see [7]). We remark that Thm. 45 is not trivial, because there exist non-extensional sensible  $\lambda$ -theories that strictly include  $\mathcal{B}$  (see [4, Exercise 16.5.5]).

Berline [7] asked whether there is a non-syntactic sensible model of lambda calculus whose theory is strictly included in  $\mathcal{B}$ . The answer is positive as shown in the following corollary.

**Theorem 46** *There exists a continuum of different sensible graph theories strictly included in  $\mathcal{B}$ .*

**Proof.** Based on a syntactic difficult result (conjectured by Kerth [23] and proved by David [17]), Kerth [23] has shown that there exists a continuum of sensible graph theories. Then the conclusion follows from Thm. 45. ■



It is well known that the  $\lambda$ -term  $\Omega$  is easy, that is, it can be consistently equated to every other closed  $\lambda$ -term  $M$ . We denote by  $(\Omega = M)^+$  the  $\lambda$ -theory generated by the equation  $\Omega = M$ .

**Theorem 47** *Let  $M$  be an arbitrary closed  $\lambda$ -term. Then we have:*

$$P =_{\mathcal{H}^*} Q, P \neq_{\mathcal{B}} Q \Rightarrow (\Omega = M)^+ \not\models P = Q.$$

*In other words,  $(\Omega = M)^+ \cap \mathcal{H}^* \subseteq \mathcal{B}$ .*

**Proof.** By [3] the  $\lambda$ -theory  $(\Omega = M)^+$  is contained within a graph theory. Then the conclusion follows from Thm. 44. ■

## 6.1 The proof of the main theorem

In this section we provide the proof of Thm. 44.

We recall that a node of a tree is a sequence of natural numbers and that the level of a node is the length of the sequence. The empty sequence will be denoted by  $\varepsilon$ .

Let  $M, N$  be closed  $\lambda$ -terms such that  $M =_{\mathcal{H}^*} N$  and  $M \neq_{\mathcal{B}} N$ . This last condition expresses the fact that the Böhm tree  $BT(M)$  of  $M$  is different from the corresponding Böhm tree  $BT(N)$  of  $N$ .

Let us give an informal overview of the proof. We start by picking a node  $u = r_1 \dots r_k$  satisfying the following two conditions: (1) the labels of  $u$  in  $BT(M)$  and  $BT(N)$  are different; (2) the labels of every strict prefix  $w = r_1 \dots r_j$  ( $j < k$ ) of  $u$  in  $BT(M)$  and  $BT(N)$  are equal. Then we show that the subterms of  $M$  and  $N$ , whose Böhm trees are the subtrees of  $BT(M)$  and  $BT(N)$  at root  $u$ , respectively, get different interpretations in all graph models. This is done in Lem. 54. In order to get the conclusion, we have to show that in all graph models it is possible to propagate upward, towards the roots of  $BT(M)$  and  $BT(N)$ , the difference “created” at node  $u$ . This is done in Lem. 55.

Let us introduce now some notations and definitions needed in the proof.

Let  $u = r_1 \dots r_k$  be a node at least level, where the labels of  $BT(M)$  and  $BT(N)$  are different. The sequence  $\varepsilon, r_1, r_1 r_2, r_1 r_2 r_3, \dots, r_1 \dots r_k$  is the sequence of nodes that are in the path from the root  $\varepsilon$  to  $u$ . These nodes will be denoted by  $u_0, u_1, u_2, \dots, u_k$ . Then, for example,  $u_0 = \varepsilon, u_2 = r_1 r_2$  and  $u_k = u$ . From the hypothesis of minimality of  $u$  it follows that

- (i) The label of the node  $u_j$  ( $0 \leq j < k$ ) in the Böhm tree of  $M$  is equal to the corresponding one in the Böhm tree of  $N$ ;
- (ii) The labels of the node  $u$  in  $BT(M)$  and  $BT(N)$  are different.

From the hypothesis  $M =_{\mathcal{H}^*} N$  and  $M \neq_{\mathcal{B}} N$  it follows that

- (iii) The node  $u$  is a starting point for a possibly infinite  $\eta$ -expansion in either  $BT(M)$  or  $BT(N)$ , but not in both. Without loss of generality, we assume to have the  $\eta$ -expansion in  $BT(N)$ .

We define two sequences  $M_{u_j}$  and  $N_{u_j}$  ( $0 \leq j \leq k$ ) of  $\lambda$ -terms whose Böhm trees  $BT(M_{u_j})$  and  $BT(N_{u_j})$  are the subtrees of  $BT(M)$  and  $BT(N)$  at root  $u_j$ , respectively. Let

$$M_{u_0} \equiv M; \quad N_{u_0} \equiv N.$$

If  $k = 0$  we have finished. Otherwise, assume by induction hypothesis that we have already defined two  $\lambda$ -terms  $M_{u_j}$  and  $N_{u_j}$  ( $j < k$ ) and that the Böhm trees of  $M_{u_j}$  and  $N_{u_j}$  are respectively the subtrees of  $BT(M)$  and  $BT(N)$  at root  $u_j$ . Assume that the principal head normal forms (principal hnfs, for short) of  $M_{u_j}$  and  $N_{u_j}$  (see [4, Def. 8.3.20]) are respectively

$$M_{u_j} =_{\lambda\beta} \lambda x_1^j \dots x_{n_j}^j \cdot z_j M_1^j \dots M_{s_j}^j; \quad (12)$$

$$N_{u_j} =_{\lambda\beta} \lambda x_1^j \dots x_{n_j}^j \cdot z_j N_1^j \dots N_{s_j}^j.$$

To abbreviate the notation we will write  $M_{u_j}$  and  $N_{u_j}$  as follows:

$$M_{u_j} =_{\lambda\beta} \lambda \bar{x}_{n_j}^j \cdot z_j M_1^j \dots M_{s_j}^j; \quad N_{u_j} =_{\lambda\beta} \lambda \bar{x}_{n_j}^j \cdot z_j N_1^j \dots N_{s_j}^j.$$

Then the node  $u_j$  in the Böhm trees of  $M$  and  $N$  has  $s_j$  sons. Since  $u_{j+1} = u_j r_{j+1}$  is a son of  $u_j$  in the Böhm trees of  $M$  and  $N$ , then we have  $r_{j+1} \leq s_j$  and we define

$$M_{u_{j+1}} \equiv M_{r_{j+1}}^j; \quad N_{u_{j+1}} \equiv N_{r_{j+1}}^j.$$

Then the Böhm trees of  $M_{u_{j+1}}$  and  $N_{u_{j+1}}$  are respectively the subtrees of  $BT(M)$  and  $BT(N)$  at root  $u_{j+1}$ . When we calculate the principal hnfs of  $M_{u_k}$  and  $N_{u_k}$  (recall that  $u_k = u$  is the node where the Böhm trees are different), we get

$$M_{u_k} \equiv M_{r_k}^{k-1} =_{\lambda\beta} \lambda \bar{x}_{n_k}^k \cdot z_k M_1^k \dots M_{s_k}^k; \quad (13)$$

$$N_{u_k} \equiv N_{r_k}^{k-1} =_{\lambda\beta} \lambda \bar{x}_{n_k}^k \cdot \lambda \bar{y}_r \cdot z_k N_1^k \dots N_{s_k}^k Q_1 \dots Q_r, \quad (14)$$

where  $y_i \leq_\eta Q_i$  ( $1 \leq i \leq r$ ) (i.e.,  $Q_i$  is a possibly infinite  $\eta$ -expansion of the variable  $y_i$ ),  $y_i$  does occur neither free nor bound in  $N_j^k$  ( $1 \leq j \leq s_k$ ) and  $Q_j$  ( $1 \leq j \neq i \leq r$ ), and it is distinct from each variable  $x_1^k, \dots, x_{n_k}^k, z_k, y_1, \dots, y_{i-1}, y_{i+1}, \dots, y_r$ .

Let  $(D, p)$  be an arbitrary graph model. First we will show that the terms  $N_{u_k}$  and  $M_{u_k}$  have different interpretations in  $(D, p)$ , that is, there exist an element  $\alpha_k \in D$  and a  $D$ -environment  $\sigma_k$  such that  $\alpha_k \in (N_{u_k})_{\sigma_k}^p$ , while  $\alpha_k \notin (M_{u_k})_{\sigma_k}^p$ . Second we will show that this difference at level  $k$  can be propagated upward, that is, there exist elements  $\alpha_i \in D$  and  $D$ -environments  $\sigma_i$  ( $i = 1, \dots, k$ ) such that  $\alpha_k \in (N_{u_k})_{\sigma_k}^p$  iff  $\alpha_i \in (N_{u_i})_{\sigma_i}^p$  iff  $\alpha_0 \in N_{\sigma_0}^p$ , and  $\alpha_k \in (M_{u_k})_{\sigma_k}^p$  iff  $\alpha_i \in (M_{u_i})_{\sigma_i}^p$  iff  $\alpha_0 \in M_{\sigma_0}^p$ .

To prove these properties of separability, we have to define the elements  $\alpha_i$  and the  $D$ -environments  $\sigma_i$ . The definition of  $\sigma_i$  is difficult and technical.

We are going to use families of points of the graph models, which are not only pairwise distinct, but also “functionally incompatible”, in the sense expressed by the following definition. Then, in the next lemma we show that such families actually exist in all graph models.

**Definition 48** *Let  $q > 1$  be a natural number. A sequence  $(\beta_n \in D : n \geq 0)$  of distinct elements of  $D$  is called a  $q$ -sequence if the following condition holds:*

$$(\forall i, j)(\forall 0 < t < q)(\forall \bar{a} \in (D^*)^t) \beta_j \neq \bar{a}_t \rightarrow \beta_i. \quad (15)$$

Recall that, if  $\bar{a} \equiv a_1 \dots a_t$ , then  $\bar{a}_t \rightarrow \beta_i$  means  $a_1 \rightarrow a_2 \rightarrow \dots \rightarrow a_t \rightarrow \beta_i$ . Notice that  $i$  may be equal to  $j$  in the above condition (15).

**Lemma 49**  *$q$ -sequences exist for every  $q > 1$ .*

**Proof.** Let  $(D, p)$  be a graph model and  $q$  be an integer greater than 1. We show that there exists a  $q$ -sequence in  $(D, p)$ .

Given  $\alpha \in D$ , we define the *degree* of  $\alpha$  as the least natural number  $k > 0$  such that there exist finite subsets  $b_1, \dots, b_k$  of  $D$  satisfying  $\alpha = b_1 \rightarrow \dots \rightarrow b_k \rightarrow \alpha$ . If such a natural number does not exist, we say that the degree of  $\alpha$  is infinite. The degree of  $\alpha$  will be denoted by  $\deg(\alpha)$ .

The proof is divided into claims.

**Claim 50** *There exists an element of  $D$  whose degree is greater than  $q$ .*

If  $D$  has an element whose degree is infinite, we are done. Otherwise, let  $\alpha_0$  be an element of  $D$  such that

$$(\forall n > 0) \alpha_0 \not\rightarrow_n \alpha_0. \quad (16)$$

Such an element does exist since otherwise the function  $p : D^* \times D \rightarrow D$  would not be total.

Let  $\alpha_i = \emptyset \rightarrow \alpha_{i-1}$  ( $i > 0$ ). In other words,  $\alpha_i = \rightarrow_i \rightarrow \alpha_0$ . We are going to show that there exists  $k$  such that  $\deg(\alpha_k) > q$ . First remark that, for all  $j$ ,  $\deg(\alpha_j) \leq \deg(\alpha_{j+1})$ , since if  $\alpha_{j+1} = b_1 \rightarrow \dots \rightarrow b_k \rightarrow \alpha_{j+1}$  then  $\alpha_j = b_2 \rightarrow \dots \rightarrow b_k \rightarrow \emptyset \rightarrow \alpha_j$ . Hence, either there exist  $j$  such that  $\deg(\alpha_j) > q$ , and we are done, or there exist  $j_0$  and  $n$  such that  $n \leq q$  and  $\deg(\alpha_j) = n$  for all  $j \geq j_0$ . We are going to show that this latter case is in fact impossible, hence concluding the proof. If  $j_0$  and  $n$  are as above, then there exist  $c_1, \dots, c_n \subset D$  such that  $\alpha_{j_0+n} = c_1 \rightarrow \dots \rightarrow c_n \rightarrow \alpha_{j_0+n}$ , i.e.  $\rightarrow_{j_0+n} \rightarrow \alpha_0 = c_1 \rightarrow \dots \rightarrow c_n \rightarrow \rightarrow_{j_0+n} \rightarrow \alpha_0$  hence  $\alpha_0 = \rightarrow_n \rightarrow \alpha_0$ , that contradicts (16).

**Claim 51** *There exists a  $q$ -sequence.*

By the above claim there exists an element  $\alpha \in D$  whose degree is greater than  $q$ . Given a family  $\{a_n\}_{n \in \omega}$  of pairwise distinct, finite subsets of  $D$ , define  $\beta_n = a_n \rightarrow \alpha$  ( $n \geq 0$ ). We

prove that the sequence  $(\beta_n : n \geq 0)$  is a  $q$ -sequence. By the way of contradiction, assume that  $\beta_i = b_1 \rightarrow \dots \rightarrow b_t \rightarrow \beta_j$  ( $0 < t < q$ ) for some  $i$  and  $j$ , i.e.,

$$a_i \rightarrow \alpha = b_1 \rightarrow \dots \rightarrow b_t \rightarrow a_j \rightarrow \alpha.$$

It follows that  $\alpha = b_2 \rightarrow \dots \rightarrow b_t \rightarrow a_j \rightarrow \alpha$ . We get a contradiction because the degree of  $\alpha$  is greater than  $q$ . ■

Let  $(\beta_n : n \geq 0)$  be a  $q$ -sequence of elements of  $D$ , where

1.  $q > (\sum_{0 \leq j \leq k} n_j) + (\sum_{0 \leq j \leq k} s_j) + r + s$ ;
2.  $n_j$  is the number of external abstractions in the principal hnf of  $M_{u_j}$  (see (12) above);
3.  $s_j$  is the number of sons of the node  $u_j$  in the Böhm tree of  $M$  (see (12) above);
4.  $r \geq 1$  is the number of  $\eta$ -expansions in  $N_{u_k}$  (see (14) above);
5.  $s$  is the number of external abstractions in the principal hnf of the subterm  $Q_r$  of  $N_{u_k}$ :

$$Q_r =_{\lambda\beta} \lambda \overline{w}_s. y_r R_1 \dots R_s \ (s \geq 0). \quad (17)$$

We now define a sequence of environments  $\rho_j$  and two sequences of elements  $\delta_j, \alpha_j \in D$  ( $0 \leq j \leq k$ ). Next the environments  $\rho_j$  will be used to define  $\sigma_0$  and  $\sigma_k$ . We start by defining  $\rho_k, \delta_k$  and  $\alpha_k$ .

- (i)  $\delta_k \equiv \rightarrow_{s_k+r-1} \rightarrow \{\rightarrow_s \rightarrow \beta_{k+1}\} \rightarrow \beta_k$ ;
- (ii)  $\rho_k(z_k) = \{\delta_k\}$ , where  $z_k$  is the head variable of the principal hnfs of  $N_{u_k}$  and  $M_{u_k}$ ;
- (iii)  $\rho_k(y_r) = \{\rightarrow_s \rightarrow \beta_{k+1}, \beta_k\}$ , where  $y_r$  is the head variable of the principal hnf of  $Q_r$ ;
- (iv)  $\rho_k(x) = \emptyset$  ( $x \neq z_k, y_r$ );
- (v)  $\alpha_k \equiv \rho_k(\overline{x}_{n_k}^k) \rightarrow \rho_k(\overline{y}_r) \rightarrow \beta_k$ .

Notice that, if  $s = 0$  (i.e., there are no external abstraction in the principal hnf of  $Q_r$ ), then by definition  $\rightarrow_0 \rightarrow \beta_{k+1}$  is just  $\beta_{k+1}$ . Moreover, the notation  $\rho_k(\overline{x}_{n_k}^k) \rightarrow \rho_k(\overline{y}_r) \rightarrow \beta_k$ , used in the definition of  $\alpha_k$ , means  $\rho_k(x_1^k) \rightarrow \dots \rightarrow \rho_k(x_{n_k}^k) \rightarrow \rho_k(y_1) \rightarrow \dots \rightarrow \rho_k(y_r) \rightarrow \beta_k$ .

Assume we have defined  $\delta_{j+1}, \alpha_{j+1}$  and  $\rho_{j+1}$  ( $j < k$ ). We define  $\delta_j, \alpha_j$  and  $\rho_j$  as follows.

- (i)  $\delta_j \equiv \rightarrow_{r_j-1} \rightarrow \{\alpha_{j+1}\} \rightarrow \rightarrow_{s_j-r_j} \rightarrow \beta_j$ ;
- (ii)  $\rho_j(z_j) = \rho_{j+1}(z_j) \cup \{\delta_j\}$ , where  $z_j$  is the head variable of the principal hnfs of  $N_{u_j}$  and  $M_{u_j}$ ;
- (iii)  $\rho_j(x) = \rho_{j+1}(x)$  ( $x \neq z_j$ );

$$(iv) \alpha_j \equiv \rho_j(\bar{x}_{n_j}^j) \rightarrow \beta_j.$$

As a matter of notation, if  $\tau$  and  $\rho$  are environments, we write  $\tau \leq \rho$  for  $\tau(x) \subseteq \rho(x)$  for all variables  $x$ .

**Lemma 52** (a)  $\rho_j \geq \rho_{j+1}$  ( $0 \leq j < k$ ).

(b) Let  $j < k$  and  $\alpha \equiv \bar{c}_t \rightarrow \beta_j$  for some sequence  $\bar{c}_t$  of length  $t < q$ . Then,  $\alpha \in \rho_0(z_j)$  iff  $\alpha \equiv \delta_j$ .

**Proof.** (a) trivially follows from the definition of  $\rho_j$ . (b) By definition of  $\rho_0$  we have that  $\gamma \in \rho_0(x)$  for some variable  $x$  iff  $\gamma$  is one of the following elements of  $D$ :  $\delta_0, \dots, \delta_k, \beta_k, \rightarrow_s \rightarrow \beta_{k+1}$ . To get the conclusion it is sufficient to apply the definition of  $q$ -sequence. ■

As a matter of notation, for every environment  $\tau$ , we write

$$\tau[\bar{x}_{n_j}^j := \rho_j(\bar{x}_{n_j}^j)] \quad (18)$$

for

$$\tau[x_1^j := \rho_j(x_1^j)] \dots [x_{n_j}^j := \rho_j(x_{n_j}^j)].$$

We now define a sequence  $\sigma_0, \dots, \sigma_{k+1}$  of environments as follows:

$$\sigma_0 = \rho_0; \quad \sigma_{j+1} = \sigma_j[\bar{x}_{n_j}^j := \rho_j(\bar{x}_{n_j}^j)] \quad (0 \leq j \leq k). \quad (19)$$

**Lemma 53** (a)  $\rho_j \leq \sigma_{j+1} \leq \rho_0$  for every  $0 \leq j \leq k$  (in particular,  $\sigma_1 = \rho_0$ ).

(b)  $\delta_j \in \sigma_{j+1}(z_j)$  for all  $0 \leq j \leq k$ .

**Proof.** (a) By definition we have  $\sigma_1 = \rho_0$ . Assume by induction hypothesis that  $\rho_{j-1} \leq \sigma_j$ . We have to show that  $\rho_j \leq \sigma_{j+1}$ . By definition  $\sigma_{j+1}(x_t^j) = \rho_j(x_t^j)$ , for every  $1 \leq t \leq n_j$ . If  $z$  is a variable distinct from  $x_t^j$  ( $1 \leq t \leq n_j$ ), then we have  $\sigma_{j+1}(z) = \sigma_j(z) \supseteq \rho_{j-1}(z) \supseteq \rho_j(z)$ , by induction hypothesis and by  $\rho_j \leq \rho_{j-1}$  (see Lem. 52).

(b) By definition  $\delta_j \in \rho_j(z_j)$ . Then the conclusion follows from  $\rho_j \leq \sigma_{j+1}$  (see (a)). ■

Finally, in the following lemma we show that  $N_{u_k}$  and  $M_{u_k}$  have different interpretations.

**Lemma 54** We have  $\alpha_k \in (N_{u_k})_{\sigma_k}^p$  and  $\alpha_k \notin (M_{u_k})_{\sigma_k}^p$ .

**Proof.** Recall that

1.  $M_{u_k} \equiv \lambda \bar{x}_{n_k}^k . z_k M_1^k \dots M_{s_k}^k$ ;
2.  $N_{u_k} \equiv \lambda \bar{x}_{n_k}^k \lambda \bar{y}_r . z_k N_1^k \dots N_{s_k}^k Q_1 \dots Q_r$ ;

3.  $Q_r \equiv \lambda \overline{w}_s . y_r R_1 \dots R_s$ ;
4.  $\delta_k \equiv \rightarrow_{s_k+r-1} \rightarrow \{ \rightarrow_s \rightarrow \beta_{k+1} \} \rightarrow \beta_k$ ;
5.  $\alpha_k \equiv \rho_k(\overline{x}_{n_k}^k) \rightarrow \rho_k(\overline{y}_r) \rightarrow \beta_k$ .

As a matter of notation, let

- $\tau \equiv \sigma_k[\overline{x}_{n_k}^k := \rho_k(\overline{x}_{n_k}^k)][\overline{y}_r := \rho_k(\overline{y}_r)]$ ;
- $\overline{Q} \equiv Q_1 \dots Q_r$ ;
- $\overline{M} \equiv M_1^k \dots M_{s_k}^k$ .
- $\overline{N} \equiv N_1^k \dots N_{s_k}^k$ .
- $\rightarrow \equiv R_1 \dots R_s$ .

By the definition of  $\sigma_{k+1}$  we immediately get that  $\tau = \sigma_{k+1}[\overline{y}_r := \rho_k(\overline{y}_r)]$ . Then we have:

$$\begin{aligned}
\alpha_k \in (N_{u_k})_{\sigma_k}^p & \text{ iff } \beta_k \in (z_k)_\tau^p \overline{N}_\tau^p \overline{Q}_\tau^p \\
& \text{ iff } \beta_k \in (z_k)_{\sigma_{k+1}}^p \overline{N}_{\sigma_{k+1}}^p \overline{Q}_\tau^p, \\
& \quad \text{by } y_i \neq z_k \text{ not free in } N_j^k \text{ and def. } \tau \\
& \text{ iff } \beta_k \in \{\delta_k\} \overline{N}_{\sigma_{k+1}}^p \overline{Q}_\tau^p, \\
& \quad \text{by } \sigma_{k+1} \leq \rho_0 \text{ and Lem. 52(b)} \\
& \text{ iff } \beta_k \in \{\delta_k\} \rightarrow_{s_k+r-1} (Q_r)_\tau^p, \\
& \quad \text{by def. } \delta_k \\
& \text{ iff } \rightarrow_s \rightarrow \beta_{k+1} \in (Q_r)_\tau^p.
\end{aligned}$$

Finally, we have:

$$\begin{aligned}
(Q_r)_\tau^p & = (\lambda \overline{w}_s . y_r R_1 \dots R_s)_\tau^p, \\
& \quad \text{by def. } Q_r \text{ (see (17) above)} \\
& = (\lambda \overline{w}_s . y_r \rightarrow)_\tau^p, \\
& \quad \text{by def. } \rightarrow \\
& = \{ \overline{c}_s \rightarrow \sigma : \sigma \in \tau(y_r) \rightarrow_{\tau[\overline{w}_s := \overline{c}_s]}^p \}, \\
& \quad \text{by } y_r \neq w_i \text{ (} i = 1, \dots, s \text{)} \\
& = \{ \overline{c}_s \rightarrow \sigma : \sigma \in \rho_k(y_r) \rightarrow_{\tau[\overline{w}_s := \overline{c}_s]}^p \}, \\
& \quad \text{by } \tau(y_r) = \rho_k(y_r) \\
& = \{ \overline{c}_s \rightarrow \sigma : \sigma \in \{ \rightarrow_s \rightarrow \beta_{k+1}, \beta_k \} \rightarrow_{\tau[\overline{w}_s := \overline{c}_s]}^p \}, \\
& \quad \text{by definition of } \rho_k(y_r) \\
& \supseteq \{ \overline{c}_s \rightarrow \sigma : \sigma \in \{ \rightarrow_s \rightarrow \beta_{k+1} \} \rightarrow_{\tau[\overline{w}_s := \overline{c}_s]}^p \} \\
& = \{ \overline{c}_s \rightarrow \beta_{k+1} : \overline{c}_s \in D^s \}.
\end{aligned}$$

Hence  $\alpha_k \in (N_{u_k})_{\sigma_k}^p$ , because  $\rightarrow_s \rightarrow \beta_{k+1} \in (Q_r)_\tau^p$ .

Recall that by (19)  $\sigma_{k+1} = \sigma_k[\bar{x}^k := \rho_k(\bar{x}^k)]$ .

$$\begin{aligned}
\alpha_k \in (M_{u_k})_{\sigma_k}^p & \text{ iff } \rho_k(\bar{y}_r) \rightarrow \beta_k \in (z_k)_{\sigma_{k+1}}^p (\bar{M})_{\sigma_{k+1}}^p \\
& \text{ iff } \rho_k(\bar{y}_r) \rightarrow \beta_k \in \{\delta_k\}(\bar{M})_{\sigma_{k+1}}^p, \\
& \quad \text{by } \sigma_{k+1} \leq \rho_0 \text{ and Lem. 52(b)} \\
& \text{ iff } \rho_k(\bar{y}_r) \rightarrow \beta_k \in \{\delta_k\} \rightarrow_{s_k}, \\
& \quad \text{by def. } \delta_k \\
& \text{ iff } \rho_k(\bar{y}_r) \rightarrow \beta_k \rightarrow_{r-1} \rightarrow \{\rightarrow_s \rightarrow \beta_{k+1}\} \rightarrow \beta_k \\
& \quad \text{by def. } \delta_k \\
& \text{ iff } \rho_k(y_r) = \{\rightarrow_s \rightarrow \beta_{k+1}\}, \\
& \quad \text{by def. } \rho_k \\
& \text{ iff } \{\rightarrow_s \rightarrow \beta_{k+1}, \beta_k\} = \{\rightarrow_s \rightarrow \beta_{k+1}\}.
\end{aligned}$$

This last relation is false. Hence  $\alpha_k \notin (M_{u_k})_{\sigma_k}^p$ . ■

The different interpretation of  $N_{u_k}$  and  $M_{u_k}$  can be propagated upward as shown in the following lemma.

**Lemma 55** *For every  $k > j \geq 0$  we have*

$$\alpha_j \in (N_{u_j})_{\sigma_j}^p \Leftrightarrow \alpha_{j+1} \in (N_{u_{j+1}})_{\sigma_{j+1}}^p$$

and

$$\alpha_j \in (M_{u_j})_{\sigma_j}^p \Leftrightarrow \alpha_{j+1} \in (M_{u_{j+1}})_{\sigma_{j+1}}^p.$$

**Proof.** We prove the result for  $N_{u_j}$ . The corresponding proof for  $M_{u_j}$  is left to the reader. We recall that  $N_{u_j} =_{\lambda\beta} \lambda \bar{x}_{n_j}^j . z_j N_1^j \dots N_{s_j}^j$ ,  $N_{u_{j+1}} \equiv N_{r_j}^j$  and  $\alpha_j \equiv \rho_j(\bar{x}_{n_j}^j) \rightarrow \beta_j$ . In the following we will write  $\bar{N}$  for  $N_1^j \dots N_{s_j}^j$ , and  $\sigma_j[\bar{x} := \rho_j(\bar{x}^j)]$  for  $\sigma_j[\bar{x}_{n_j}^j := \rho_j(\bar{x}_{n_j}^j)]$ .

$$\begin{aligned}
\alpha_j \in (N_{u_j})_{\sigma_j}^p & \text{ iff } \beta_j \in (z_j)_{\sigma_j[\bar{x}:=\rho_j(\bar{x}^j)]}^p \bar{N}_{\sigma_j[\bar{x}:=\rho_j(\bar{x}^j)]}^p \\
& \quad \text{by def. } \alpha_j \\
& \text{ iff } \beta_j \in (z_j)_{\sigma_{j+1}}^p (\bar{N}^j)_{\sigma_{j+1}}^p, \\
& \quad \text{by def. } \sigma_{j+1} \\
& \text{ iff } \beta_j \in \{\delta_j\}(\bar{N}^j)_{\sigma_{j+1}}^p, \\
& \quad \text{by } \sigma_{j+1} \leq \rho_0, \text{ Lem.52(b), 53(b)} \\
& \text{ iff } \beta_j \in \{\delta_j\} \rightarrow^{r_j-1} (N_{r_j}^j)_{\sigma_{j+1}}^p \rightarrow^{s_j-r_j}, \\
& \quad \text{by def. } \delta_j \\
& \text{ iff } \alpha_{j+1} \in (N_{u_{j+1}})_{\sigma_{j+1}}^p, \\
& \quad \text{by } N_{u_{j+1}} \equiv N_{r_j}^j \text{ and def. } \delta_j.
\end{aligned}$$

The conclusion of the lemma is now immediate. ■

**Lemma 56** *We have  $\alpha_0 \in N_{\sigma_0}^p$ , while  $\alpha_0 \notin M_{\sigma_0}^p$ .*

**Proof.** Recall that  $N \equiv N_{u_0}$  and  $M \equiv M_{u_0}$ . By applying Lem. 55 it is easy to show that that  $\alpha_0 \in N_{\sigma_0}^p \Leftrightarrow \alpha_k \in (N_{u_k})_{\sigma_k}^p$ , and  $\alpha_0 \in M_{\sigma_0}^p \Leftrightarrow \alpha_k \in (M_{u_k})_{\sigma_k}^p$ . Then the conclusion is immediate, because by Lem. 54 we have that  $\alpha_k \in (N_{u_k})_{\sigma_k}^p$  and  $\alpha_k \notin (M_{u_k})_{\sigma_k}^p$ . ■

## 7 Conclusion and future work

In this paper, we have collected in an organized manner several already published results and some new material: the existence of the minimum (resp. minimum sensible) graph-theory appeared originally in [12] (resp. [13]). The new presentation of section 3 stresses the relevance and generality of the *weak product* construction, underlying these results, and add some new results (for instance, the fact that the theory of weak products is semisensible and it is in general strictly finer than the intersection of the factors' theories, obtained via the notion of *self weak product*).

Section 6 covers the main result of [13], namely the fact that the maximal sensible graph theory is  $\mathcal{B}$ .

The content of Section 5, a negative answer to the question of whether  $\lambda\beta$  is the minimal graph theory, also appeared in [13]. Actually, this negative result opens the way to the investigation of the minimal graph theory.

Section 3.2 and Section 3.3 present new results. First, the fact that stratified graph models, which are those obtained by canonical completion of partial pairs (i.e. virtually all known graph models, apart from those constructed by *forcing* [3, 8]) have semisensible theories. Then we show that the theory of a weak product is in general strictly finer than the intersection of the factors' theories. Finally, we provide equations between unsolvable terms which are not preserved in weak products.

Several questions remains open. Among them, we wish to address that concerning the minimal sensible graph theory: Is it  $\mathcal{H}$  (the minimum sensible theory) or is it bigger? For the time being, we are able to separate in a graph model some typical example of  $\mathcal{B}$ -equivalent,  $\mathcal{H}$ -distinct  $\lambda$ -terms, like  $Yx$  and  $\Theta x$ .

The notion of *effective* graph model is a natural one: it is enough to ask that the coding function be total recursive w.r.t. given enumerations of the model's web, finite sets and pairs of natural numbers. Then one recasts classical recursion theory results in the framework of graph models, and this seems particularly compelling since those are models of the  $\lambda$ -calculus. A forthcoming paper on this subject is [9].



## References

- [1] S. Abramsky. Domain theory in logical form. *Annals of Pure and Applied Logic*, 51:1–77, 1991.
- [2] F. Alessi, M. Dezani, F. Honsell. Filter models and easy terms. ICTCS’01, LNCS 2202, Springer-Verlag, 17–37, 2001.
- [3] J. Baeten, B. Boerboom. Omega can be anything it should not be. *Indag. Mathematicae*, 41:111–120, 1979.
- [4] H.P. Barendregt. *The lambda calculus: Its syntax and semantics*. North-Holland Publishing Co., Amsterdam, 1984.
- [5] H.P. Barendregt, M. Coppo, M. Dezani-Ciancaglini. A filter lambda model and the completeness of type assignment. *J. Symbolic Logic*, 48:931–940, 1984.
- [6] O. Bastonero, X. Gouy. Strong stability and the incompleteness of stable models of  $\lambda$ -calculus. *Annals of Pure and Applied Logic*, 100:247–277, 1999.
- [7] C. Berline. From computation to foundations via functions and application: The  $\lambda$ -calculus and its webbed models. *Theoretical Computer Science*, 249:81–161, 2000.
- [8] C. Berline, A. Salibra. Easiness in graph models, *Theoretical Computer Science* 354:4–23, 2006.
- [9] C. Berline, G. Manzonetto, A. Salibra. Recursively enumerable  $\lambda$ -theories and Visser topology. *In preparation*.
- [10] G. Berry. Stable models of typed lambda-calculi. In *Proc. 5th Int. Coll. on Automata, Languages and Programming*, LNCS 62, Springer-Verlag, Berline, 1978.
- [11] A. Bucciarelli, T. Ehrhard. Sequentiality and strong stability. *Sixth Annual IEEE Symposium on Logic in Computer Science*, IEEE Computer Society Press, 138–145, 1991.
- [12] A. Bucciarelli, A. Salibra. The minimal graph model of lambda calculus. *28th International Symposium on Mathematical Foundations of Computer Science*, LNCS 2747, Springer-Verlag, 2003.
- [13] A. Bucciarelli, A. Salibra. The sensible graph theories of lambda calculus. *19th Annual IEEE Symposium on Logic in Computer Science (LICS 2004)*, IEEE Computer Society Press, 2004.
- [14] A. Church. A set of postulates for the foundation of logic. *Annals of Math.*, 2:346–366, 1933.
- [15] A. Church. *The calculi of lambda conversion*. Princeton University Press, Princeton, 1941.

- [16] M. Coppo, M. Dezani. An extension of the basic functionality theory for the  $\lambda$ -calculus. *Notre Dame J. Formal Logic*, 21:685–693, 1980.
- [17] R. David. Computing with Böhm trees. *Fundamenta Informaticae*, 45:53–77, 2001.
- [18] Di Gianantonio, P., Honsell, F.: An abstract notion of application, Typed lambda calculi and applications, LNCS 664, Springer-Verlag, 124–138, 1993.
- [19] P. Di Gianantonio, F. Honsell, G.D. Plotkin. Uncountable limits and the lambda calculus. *Nordic J. Comput.*, 2:126–145, 1995.
- [20] X. Gouy. Etude des théories équationnelles et des propriétés algébriques des modèles stables du  $\lambda$ -calcul. Thèse, Université de Paris 7, 1995.
- [21] F. Honsell, S. Ronchi della Rocca. An approximation theorem for topological lambda models and the topological incompleteness of lambda calculus. *Journal of Computer and System Sciences* 45:49–75, 1992.
- [22] R. Kerth. *Isomorphisme et équivalence équationnelle entre modèles du  $\lambda$ -calcul*. Thèse, Université de Paris 7, 1995.
- [23] R. Kerth. Isomorphism and equational equivalence of continuous lambda models. *Studia Logica* 61:403–415, 1998.
- [24] R. Kerth. On the construction of stable models of  $\lambda$ -calculus. *Theoretical Computer Science* 269:23–46, 2001.
- [25] G. Longo. Set-theoretical models of  $\lambda$ -calculus: theories, expansions and isomorphisms. *Ann. Pure Applied Logic* 24:153–188, 1983.
- [26] G.D. Plotkin. Set-theoretical and other elementary models of the  $\lambda$ -calculus. *Theoretical Computer Science* 121:351–409, 1993.
- [27] A. Salibra. A continuum of theories of lambda calculus without semantics. *16th Annual IEEE Symposium on Logic in Computer Science*, IEEE Computer Society Press, 334–343, 2001.
- [28] A. Salibra. Topological incompleteness and order incompleteness of the lambda calculus. (LICS'01 Special Issue). *ACM Transactions on Computational Logic* 4:379–401, 2003.
- [29] D.S. Scott. Continuous lattices. In *Toposes, Algebraic geometry and Logic*. LNM 274, Springer-Verlag, Berlin, 1972.
- [30] P. Selinger. Order-incompleteness and finite lambda reduction models. *Theoretical Computer Science* 309:43–63, 2003.

Antonio BUCCIARELLI

Université Paris 7

Equipe PPS (case 7014)

2, place Jussieu

72251 Paris Cedex 05, France.

e-mail : buccia@pps.jussieu.fr

<http://www.pps.jussieu.fr/~buccia>

\*\*\*\*\*

Antonino SALIBRA

Dipartimento di Informatica,

Università Ca' Foscari di Venezia

Via Torino 155

30172 Venezia, Italy

e-mail: salibra@dsi.unive.it

<http://www.dsi.unive.it/~salibra>

\*\*\*\*\*